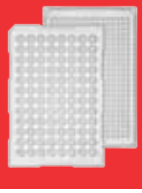
applied biosystems

Real-time PCR handbook

ThermoFisher SCIENTIFIC

Single-tube assays



96- and 384-well plates

384-well TaqMan Array Cards

OpenArray plates

Commonly used formats for real-time PCR.

Basics of real-time PCR

1

Contents

1.1 Introduction

The polymerase chain reaction (PCR) is one of the most powerful technologies in molecular biology. Using PCR, specific sequences within a DNA or complementary DNA (cDNA) template can be copied, or "amplified," many thousand- to million-fold using sequence-specific oligonucleotides, heat-stable DNA polymerase, and thermal cycling. In traditional (endpoint) PCR, detection and quantification of the amplified sequence are performed at the end of the reaction after the last PCR cycle, and involve post-PCR analysis such as gel electrophoresis and image analysis. In real-time quantitative PCR (often shortened to real-time PCR or qPCR), PCR product is measured at each cycle. By monitoring reactions during the exponential amplification phase of the reaction, users can determine the initial quantity of the target with great precision.

PCR theoretically amplifies DNA exponentially, doubling the number of target molecules with each amplification cycle. When it was first developed, scientists reasoned that the number of cycles and the amount of PCR end-product could be used to calculate the initial quantity of genetic material by comparison with a known standard. To address the need for robust quantification, the technique of real-time PCR was developed. Currently, endpoint PCR is used mostly to amplify specific DNA for sequencing, cloning, and use in other molecular biology techniques.

In real-time PCR, the amount of DNA is measured after each cycle via fluorescent dyes that yield increasing fluorescent signal in direct proportion to the number of PCR product molecules (amplicons) generated. Data collected in the exponential phase of the reaction yield quantitative information on the starting quantity of the amplification target. Fluorescent reporters used in real-time PCR include double-stranded DNA (dsDNA)–binding dyes, or dye molecules attached to PCR primers or probes that hybridize with PCR products during amplification.

The change in fluorescence over the course of the reaction is measured by an instrument that combines thermal cycling with fluorescent dye scanning capability. By plotting fluorescence against the cycle number, the real-time PCR instrument generates an amplification plot that represents the accumulation of product over the duration of the entire PCR (Figure 1).

The advantages of real-time PCR include:

- Ability to monitor the progress of individual PCR reactions as they occur in real time
- Ability to precisely measure the amount of amplicon at each cycle, which allows highly accurate quantification of the amount of starting material in samples
- An increased dynamic range of detection
- Amplification and detection occur in a single tube, eliminating post-PCR manipulations

Over the past several years, real-time PCR has become the leading tool for the detection and quantification of DNA or RNA. Using these techniques allows precise detection that is accurate within a 2-fold range, with a dynamic range of input material covering 6 to 8 orders of magnitude.

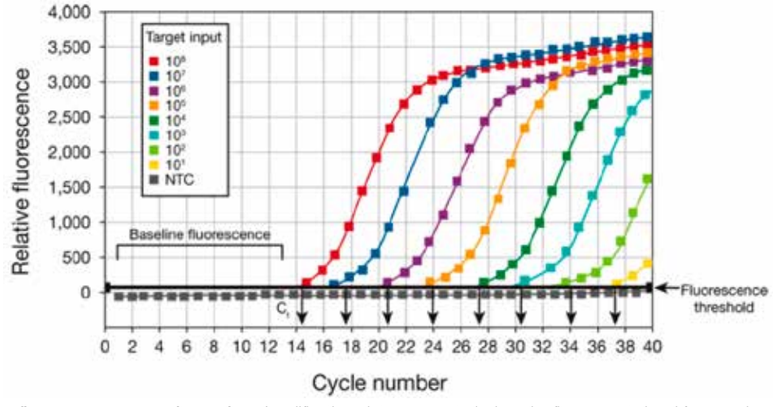


Figure 1. Relative fluorescence vs. cycle number. Amplification plots are created when the fluorescent signal from each sample is plotted against cycle number; therefore, amplification plots represent the accumulation of product over the duration of the real-time PCR experiment. The samples used to create the plots in this figure are a dilution series of the target DNA sequence. NTC, no-template control.

1.2 Overview of real-time PCR

This section provides an overview of the steps involved in performing real-time PCR. Real-time PCR is a variation of the standard PCR technique that is commonly used to quantify DNA or RNA in a sample. Using sequence-specific primers, the number of copies of a particular DNA or RNA sequence can be determined. By measuring the amount of amplified product at each stage during the PCR cycle, quantification is possible. If a particular sequence (DNA or RNA) is abundant in the sample, amplification is observed in earlier cycles; if the sequence is scarce, amplification is observed in later cycles. Quantification of amplified product is obtained using fluorescent probes or fluorescent DNA-binding dyes and real-time PCR instruments that measure fluorescence while performing the thermal cycling for PCR.

Real-time PCR steps

There are three major steps that make up each cycle in real-time PCR. Reactions are generally run for 40 cycles.

1. **Denaturation**—high-temperature incubation is used to "melt" dsDNA into single strands and loosen secondary structure in single-stranded DNA (ssDNA). The highest temperature that the DNA polymerase can withstand is typically used (usually 95°C). The denaturation time can be increased if template GC content is high.
2. **Annealing**—during annealing, complementary sequences have an opportunity to hybridize, so an appropriate temperature is used that is based on the calculated melting temperature (T_m) of the primers (typically 5°C below the T_m of the primer).
3. **Extension**—at 70–72°C, the activity of the DNA polymerase is optimal, and primer extension occurs at rates of up to 100 bases per second. When an amplicon in real-time PCR is small, this step is often combined with the annealing step, using 60°C as the temperature.

Two-step qRT-PCR

Two-step quantitative reverse-transcription PCR (qRT-PCR) starts with the reverse transcription of either total RNA or poly(A) RNA into cDNA using a reverse transcriptase. This first-strand cDNA synthesis reaction can be primed using random primers, oligo(dT), or gene-specific primers. To give an equal representation of all targets in real-time PCR applications and to avoid the 3′ bias of oligo(dT) primers, many researchers use random primers or a mixture of oligo(dT) and random primers.

The temperature used for cDNA synthesis depends on the reverse transcriptase chosen. After reverse transcription, approximately 10% of the cDNA is transferred to a separate tube for real-time PCR.

One-step qRT-PCR

One-step qRT-PCR combines the first-strand cDNA synthesis reaction and real-time PCR reaction in the same tube, simplifying reaction setup and reducing the possibility of contamination. Gene-specific primers are required. This is because using oligo(dT) or random primers will generate nonspecific products in the one-step procedure and reduce the amount of product of interest.

1.3 Overview of real-time PCR components

This section provides an overview of the major reaction components and parameters involved in real-time PCR experiments. A more detailed discussion of specific components like reporter dyes, passive reference dyes, and uracil DNA glycosylase (UDG) is provided in subsequent sections of this handbook.

DNA polymerase

PCR performance is often related to the thermostable DNA polymerase, so enzyme selection is critical to success. One of the main factors affecting PCR specificity is the fact that *Taq* DNA polymerase has residual activity at low temperatures. Primers can anneal nonspecifically to DNA during reaction setup, allowing the polymerase to synthesize nonspecific product. The problem of nonspecific products resulting from mispriming can be minimized by using a "hot-start" enzyme. Using a hot-start enzyme ensures that DNA polymerase is not active during reaction setup and the initial DNA denaturation step.

Reverse transcriptase

The reverse transcriptase is as critical to the success of qRT-PCR as the DNA polymerase. It is important to choose a reverse transcriptase that not only provides high yields of full-length cDNA, but also has good activity at high temperatures. High-temperature performance is also very important for denaturation of RNA with secondary structure. In one-step qRT-PCR, a reverse transcriptase that retains its activity at higher temperatures allows use of a gene-specific primer with a high T_m, increasing specificity and reducing background.

dNTPs

It is a good idea to purchase both the dNTPs and the thermostable DNA polymerase from the same vendor, as it is not uncommon to see a loss in sensitivity of one full threshold cycle (C_t) in experiments that employ these reagents from separate vendors.

Magnesium concentration

In real-time PCR, magnesium chloride or magnesium sulfate is typically used at a final concentration of 3 mM. This concentration works well for most targets; however, the optimal magnesium concentration may vary between 3 and 6 mM.

Good experimental technique

Do not underestimate the importance of good laboratory technique. It is best to use dedicated equipment and solutions for each stage of the reactions, from preparation of the template to post-PCR analysis. The use of aerosol-barrier tips and screwcap tubes can help decrease cross-contamination problems. To obtain tight data from replicates (ideally, triplicates), prepare a master mix that contains all the reaction components except sample. The use of a master mix reduces the number of pipetting steps and, consequently, reduces the chances of cross-well contamination and other pipetting errors.

Template

Use 10 to 1,000 copies of template nucleic acid for each real-time PCR reaction. This is equivalent to approximately 100 pg to 1 µg of genomic DNA (gDNA), or cDNA generated from 1 pg to 100 ng of total RNA. Excess template may also bring higher contaminant levels that can greatly reduce PCR efficiency. Depending on the specificity of the PCR primers for cDNA rather than gDNA, it may be important to treat RNA templates to reduce the chance that they contain gDNA contamination. One option is to treat the template with DNase I.

Pure, intact RNA is essential for full-length, high-quality cDNA synthesis and may be important for accurate mRNA quantification. RNA should be devoid of any RNase contamination, and aseptic conditions should be maintained. Total RNA typically works well in qRT-PCR; isolation of mRNA is typically not necessary, although it may improve the yield of specific cDNAs.

Real-time PCR primer design

Good primer design is one of the most important parameters in real-time PCR. This is why many researchers choose to purchase Applied Biosystems™ TaqMan™ Assay products—primers and probes for real-time PCR designed using a proven algorithm and trusted by scientists around the world. When designing real-time PCR primers, keep in mind that the amplicon length should be approximately 50–150 bp, since longer products do not amplify as efficiently.

In general, primers should be 18–24 nucleotides in length. This provides for practical annealing temperatures. Primers should be designed according to standard PCR guidelines. They should be specific for the target sequence and be free of internal secondary structure. Primers should avoid stretches of homopolymer sequences (e.g., poly(dG)) or repeating motifs, as these can hybridize inappropriately.

Primer pairs should have compatible melting temperatures (within 1°C) and contain approximately 50% GC content. Primers with high GC content can form stable imperfect hybrids. Conversely, high AT content depresses the T_m of perfectly matched hybrids. If possible, the 3′ end of the primer should be GC-rich to enhance annealing of the end that will be extended. Analyze primer pair sequences to avoid complementarity and hybridization between primers (primer-dimers).

For qRT-PCR, design primers that anneal to exons on both sides of an intron (or span an exon/exon boundary of the mRNA) to allow differentiation between amplification of cDNA and potential contaminating gDNA by melting curve analysis. To confirm the specificity of primers, perform a BLAST™ search against public databases to be sure that they only recognize the target of interest.

Optimal results may require a titration of primer concentrations between 50 and 500 nM. A final concentration of 200 nM for each primer is effective for most reactions.

Primer design software

Primer design software programs, such as Invitrogen™ OligoPerfect™ designer and Applied Biosystems™ Primer Express™ software, in addition to sequence analysis software, such as Invitrogen™ Vector NTI™ Software, can automatically evaluate a target sequence and design primers for it based on the criteria previously discussed.

At a minimum, using primer design software will ensure that primers are specific for the target sequence and free of internal secondary structure, and avoid complementary hybridization at 3′ ends within each primer and with each other. As mentioned previously, good primer design is especially critical when using DNA-binding dyes for amplicon detection.

1.4 Real-time PCR analysis technology

This section defines the major terms used in real-time PCR analysis.

Baseline

The baseline of the real-time PCR reaction refers to the signal level during the initial cycles of PCR, usually cycles 3 to 15, in which there is little change in fluorescent signal. The low-level signal of the baseline can be equated to the background or the "noise" of the reaction (Figure 2). The baseline in real-time PCR is determined empirically for each reaction, by user analysis or automated analysis of the amplification plot. The baseline should be set carefully to allow accurate determination of the threshold cycle (C_t), defined below. The baseline determination should take into account enough cycles to eliminate the background found in the early cycles of amplification, but should not include the cycles in which the amplification signal begins to rise above background. When comparing different real-time PCR reactions or experiments, the baseline should be defined in the same way for each (Figure 2).

Threshold

The threshold of the real-time PCR reaction is the level of signal that reflects a statistically significant increase over the calculated baseline signal (Figure 2). It is set to distinguish relevant amplification signal from the background. Usually, real-time PCR instrument software automatically sets the threshold at 10 times the standard deviation of the fluorescence value of the baseline. However, the position of the threshold can be set at any point in the exponential phase of PCR.

Threshold cycle (C_t)

C_t is the cycle number at which the fluorescent signal of the reaction crosses the threshold. The C_t is used to calculate the initial DNA copy number, because the C_t value is inversely related to the starting amount of target. For example, in comparing real-time PCR results from samples containing different amounts of target, a sample with twice the starting amount will yield a C_t value one cycle earlier than a sample with half the number of copies of the target. This assumes that the PCR is operating at 100% efficiency (i.e., the amount of product doubles perfectly during each cycle) in both reactions.

As the template amount decreases, the cycle number at which significant amplification is seen increases. With a 10-fold dilution series, the C_t values are ~3.3 cycles apart (Figure 3).

Standard curve

A dilution series of known template concentrations can be used to establish a standard curve for determining the initial starting amount of the target template in experimental samples or for assessing the reaction efficiency (Figure 4). The log of each known concentration in the dilution series (x-axis) is plotted against the C_t value for that concentration (y-axis). From this standard curve, information about the performance of the reaction as well as various reaction parameters (including slope, y-intercept, and correlation coefficient) can be derived. The concentrations chosen for the standard curve should encompass the expected concentration range of the target in the experimental samples.

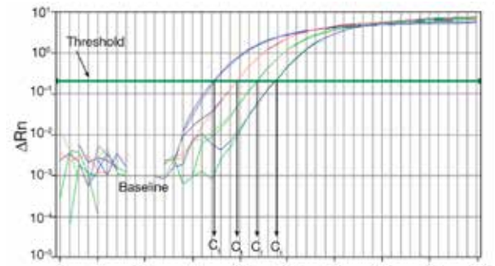


Figure 2. The baseline and threshold of a real-time PCR reaction.

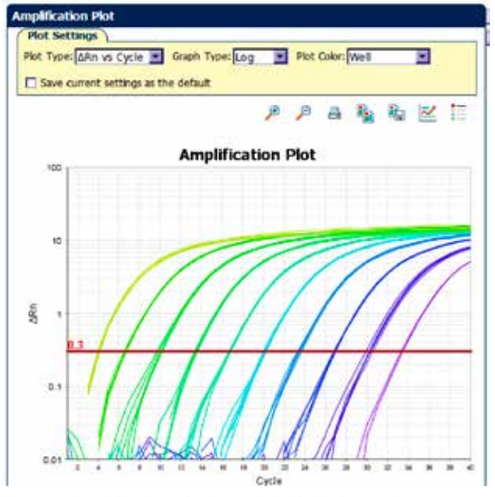


Figure 3. Amplification plot for a 10-fold dilution series.

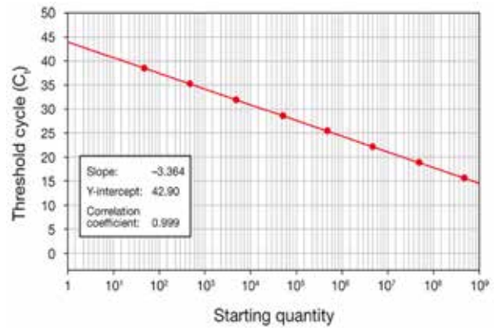


Figure 4. Example of a standard curve of real-time PCR data. A standard curve shows threshold cycle (C_t) on the y-axis and the starting quantity of RNA or DNA target on the x-axis. Slope, y-intercept, and correlation coefficient values are used to provide information about the performance of the reaction.

Correlation coefficient (R^2)

The correlation coefficient is a measure of how well the data fit the standard curve. The R^2 value reflects the linearity of the standard curve. Ideally, $R^2 = 1$, although 0.999 is generally the maximum value.

Y-intercept

The y-intercept corresponds to the theoretical limit of detection of the reaction, or the C_t value expected if the lowest copy number of target molecules denoted on the x-axis gave rise to statistically significant amplification. Though PCR is theoretically capable of detecting a single copy of a target, a copy number of 2–10 is commonly specified as the lowest target level that can be reliably quantified in real-time PCR applications. This limits the usefulness of the y-intercept value as a direct measure of sensitivity. However, the y-intercept value may be useful for comparing different amplification systems and targets.

Exponential phase

It is important to quantify real-time PCR in the early part of the exponential phase as opposed to in the later cycles or when the reaction reaches the plateau. At the beginning of the exponential phase, all reagents are still in excess, the DNA polymerase is still highly efficient, and the amplification product, which is present in a low amount, will not compete with the primers' annealing capabilities. All of these factors contribute to more accurate data.

Slope

The slope of the log-linear phase of the amplification reaction is a measure of reaction efficiency. To obtain accurate and reproducible results, reactions should have an efficiency as close to 100% as possible, equivalent to a slope of –3.32 (see Efficiency, below, for more detail).

Efficiency

A PCR efficiency of 100% corresponds to a slope of –3.32, as determined by the following equation:

$$\text{Efficiency} = 10^{(-1/\text{slope})} - 1$$

Ideally, the efficiency of PCR should be 100%, meaning the template doubles after each thermal cycle during exponential amplification. The actual efficiency can give valuable information about the reaction. Experimental factors such as the length, secondary structure, and GC content of the amplicon can influence efficiency. Other conditions that may influence efficiency are the dynamics of the reaction itself, the use of non-optimal reagent concentrations, and enzyme quality, which can result in efficiencies below 90%. The presence of PCR inhibitors in one or more of the reagents can produce efficiencies of greater than 110%. A good reaction should have an efficiency between 90% and 110%, which corresponds to a slope between –3.58 and –3.10.

Dynamic range

This is the range over which an increase in starting material concentration results in a corresponding increase in amplification product. Ideally, the dynamic range for real-time PCR should be 7–8 orders of magnitude for plasmid DNA and at least a 3–4 orders of magnitude for cDNA or gDNA.

Absolute quantification

Absolute quantification describes a real-time PCR experiment in which samples of known quantity are serially diluted and then amplified to generate a standard curve. Unknown samples are then quantified by comparison with this curve.

Relative quantification

Relative quantification describes a real-time PCR experiment in which the expression of a gene of interest in one sample (i.e., treated) is compared to expression of the same gene in another sample (i.e., untreated). The results are expressed as fold change (increase or decrease) in expression of the treated sample in relation to the untreated sample. A normalizer gene (such as β-actin) is used as a control for experimental variability in this type of quantification.

Melting curve (dissociation curve)

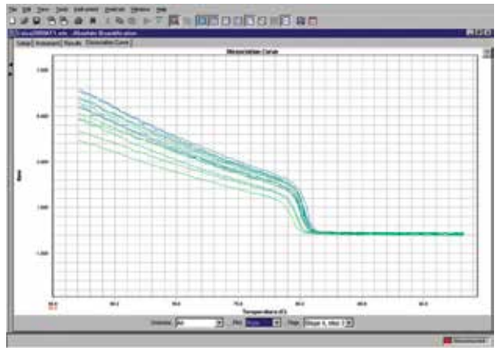
A melting curve charts the change in fluorescence observed when dsDNA with incorporated dye molecules dissociates ("melts") into single-stranded DNA (ssDNA) as the temperature of the reaction is raised. For example, when dsDNA bound with SYBR™ Green I dye is heated, a sudden decrease in fluorescence is detected when the melting temperature (T_m) is reached, due to dissociation of the DNA strands and subsequent release of the dye. The fluorescence is plotted against temperature (Figure 5A), and then the –ΔF/ΔT (change in fluorescence/change in temperature) value is plotted against temperature to obtain a clear view of the melting dynamics (Figure 5B).

Post-amplification melting-curve analysis is a simple, straightforward way to check real-time PCR reactions for primer-dimer artifacts and to ensure reaction specificity. Because the melting temperature of nucleic acids is affected by length, GC content, and the presence of base mismatches, among other factors, different PCR products can often be distinguished by their melting characteristics. The characterization of reaction products (e.g., primer-dimers vs. amplicons) via melting curve analysis reduces the need for time-consuming gel electrophoresis.

The typical real-time PCR data set shown in Figure 6 illustrates many of the terms that have been discussed. Figure 6A illustrates a typical real-time PCR amplification plot. During the early cycles of the PCR, there is little change in the fluorescent signal. As the reaction progresses, the level of fluorescence begins to increase with each cycle. The reaction threshold is set above the baseline in the exponential portion of the plot. This threshold is used to assign the C_t value of each amplification reaction. C_t values for a series of reactions containing a known quantity of target can be used to generate a standard curve. Quantification is performed by comparing C_t values for unknown samples against this standard curve or, in the case of relative quantification, against each other, with the standard curve serving as an efficiency check. C_t values are inversely related to the amount of starting template: the higher the amount of starting template in a reaction, the lower the C_t value for that reaction.

Figure 6B shows the standard curve generated from the C_t values in the amplification plot. The standard curve provides important information regarding the amplification efficiency, replicate consistency, and theoretical detection limit of the reaction.

A



B

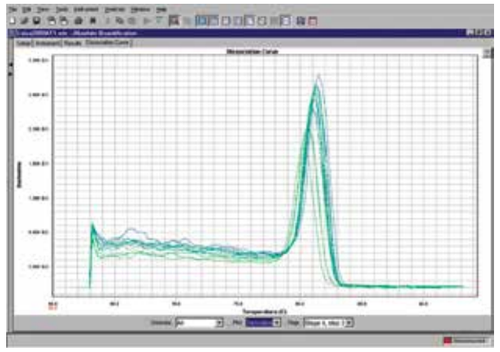
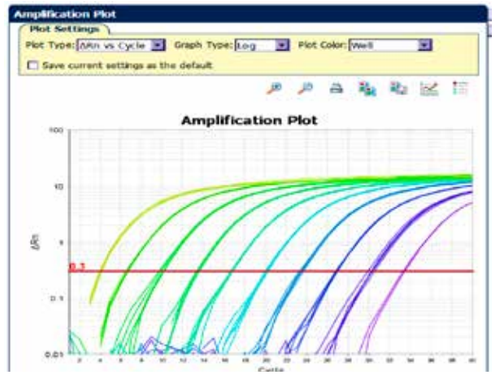


Figure 5. Melting curve analysis. (A) Melting curve and **(B)** –ΔF/ΔT vs. temperature.

A



B

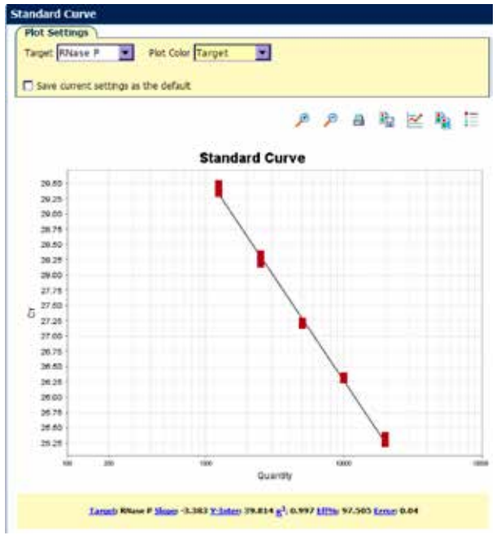


Figure 6. Amplification of RNase P DNA ranging from 1.25 x 10^3 to 2 x 10^4 copies. Real-time PCR of 2-fold serial dilutions of human RNase P DNA was performed using a FAM™ dye–labeled TaqMan Assay with Applied Biosystems™ TaqMan™ Universal Master Mix II, under standard thermal cycling conditions on an Applied Biosystems™ ViiA™ 7 Real-Time PCR System. **(A)** Amplification plot. **(B)** Standard curve showing copy number of template vs. C_t.

1.5 Real-time PCR fluorescence detection systems

Real-time fluorescent PCR chemistries

Many real-time fluorescent PCR chemistries exist, but the most widely used are 5′ nuclease assays such as TaqMan Assays and SYBR Green dye–based assays (Figure 7).

The 5′ nuclease assay is named for the 5′ nuclease activity associated with *Taq* DNA polymerase (Figure 8).

The 5′ nuclease domain has the ability to degrade DNA bound to the template, downstream of DNA synthesis. A second key element in the 5′ nuclease assay is a phenomenon called fluorescence resonance energy transfer (FRET). In FRET, the emissions of a fluorescent dye can be strongly reduced by the presence of another dye, often called the quencher, in close proximity.

FRET can be illustrated by two fluorescent dyes: green and red (Figure 9). The green fluorescent dye has a higher energy of emission compared to the red, because green light has a shorter wavelength than red. If the red dye is in close proximity to the green dye, excitation of the green dye will cause its emitted energy to be transferred to the red dye. In other words, energy is being transferred from a higher to a lower level. Consequently, the signal from the green dye will be suppressed or "quenched." However, if the two dyes are not in close proximity, FRET cannot occur, so the green dye emits light with no loss of signal.

A 5′ nuclease assay for target detection or quantification typically consists of two PCR primers and a TaqMan probe (Figure 10).

Before PCR begins, the TaqMan probe is intact and has a degree of flexibility. While the probe is intact, the reporter and quencher have a natural affinity for each other, allowing FRET to occur (Figure 11). The reporter signal is quenched prior to PCR.

During PCR, the primers and probe anneal to the target. DNA polymerase extends the primer upstream of the probe. If the probe is bound to the correct target sequence, the polymerase's 5′ nuclease activity cleaves the probe, releasing a fragment containing the reporter dye. Once cleavage takes place, the reporter and quencher dyes are no longer attracted to each other; the released reporter molecule will no longer be quenched.

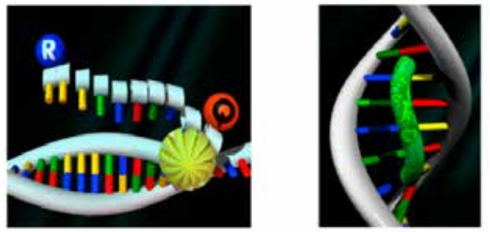


Figure 7. Representation of a 5′ nuclease assay (left) and SYBR Green dye binding to DNA (right).

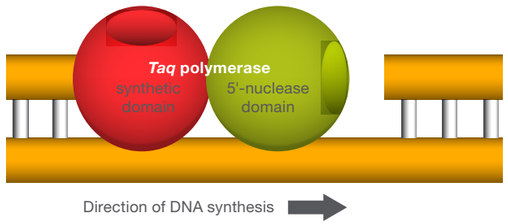


Figure 8. A representation of *Taq* DNA polymerase. Each colored sphere represents a protein domain.

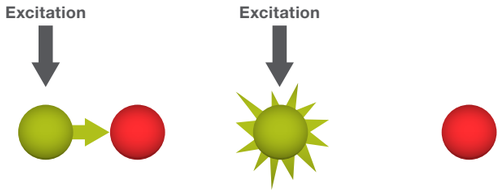


Figure 9. Example of the FRET phenomenon. (A) FRET occurs when a green light–emitting fluorescent dye is in close proximity to a red light–emitting fluorescent dye. **(B)** FRET does not occur when the two fluorescent dyes are not in close proximity.

5′ nuclease assay specificity

Assay specificity is the degree to which the assay includes signal from the target and excludes signal from non-target in the results. Specificity is arguably the most important aspect of any assay. The greatest threat to assay specificity for 5′ nuclease assays is the presence of homologs. Homologs are genes similar in sequence to that of the target, but they are not the intended target of the assay. Homologs are extremely common within species and across related species.

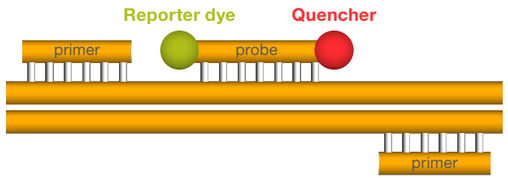


Figure 10. TaqMan probe. The TaqMan probe has a gene-specific sequence and is designed to bind the target between the two PCR primers. Attached to the 5′ end of the TaqMan probe is the "reporter," which is a fluorescent dye that will report the amplification of the target. On the 3′ end of the probe is a quencher, which quenches fluorescence from the reporter in intact probes. The quencher also blocks the 3′ end of the probe so that it cannot be extended by thermostable DNA polymerase.

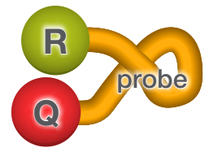


Figure 11. Representation of a TaqMan probe in solution. R is the reporter dye, Q is the quencher molecule, and the orange line represents the oligonucleotide.

5′ nuclease assays offer two tools for specificity: primers and probes. A mismatch between the target and homolog positioned at the 3′-most base of the primer has maximal impact on the specificity of the primer. A mismatch further away from the 3′ end will have less impact on specificity. In contrast, mismatches across most of the length of an Applied Biosystems™ TaqMan™ MGB probe, which is shorter than an Applied Biosystems™ TaqMan™ TAMRA™ probe, can have a strong impact on specificity—TaqMan MGB probes are stronger tools for specificity than primers.

For example, a 1- or 2-base random mismatch in the primer binding site will very likely allow the DNA polymerase to extend the primer bound to the homolog with high efficiency. A 1- or 2-base extension by DNA polymerase will stabilize the primer bound to the homolog, so it is just as stably bound as primer bound to the intended, fully complementary target. At that point, there is nothing to prevent the DNA polymerase from continuing synthesis to produce a copy of the homolog.

In contrast, mismatches on the 5′ end of the TaqMan probe binding site cannot be stabilized by the DNA polymerase due to the quencher block on the 3′ end. Mismatches in a TaqMan MGB probe binding site will reduce how tightly the probe is bound, so that instead of cleavage, the intact probe is displaced. The intact probe returns to its quenched configuration, so that when data are collected at the end of the PCR cycle, signal is produced from the target but not from the homolog, even though the homolog is being amplified.

In addition to homologs, PCR may also amplify nonspecific products produced by primers binding to seemingly random locations in the sample DNA or sometimes to themselves in so-called primer-dimers. Since nonspecific products are unrelated to the target, they do not have TaqMan probe binding sites, and thus are not seen in the real-time PCR data.

TaqMan probe types

TaqMan probes may be divided into two types: MGB and non-MGB. The first TaqMan probes could be classified as non-MGB. They used TAMRA dye as the quencher. Early in the development of real-time PCR, extensive testing revealed that TaqMan probes required an annealing temperature significantly higher than that of PCR primers to allow cleavage to take place. TaqMan probes were therefore longer than primers. A 1-base mismatch in such long probes had a relatively mild effect on probe binding, allowing cleavage to take place. However, for many applications involving high genetic complexity, such as eukaryotic gene expression and single-nucleotide polymorphisms (SNPs), a higher degree of specificity was desirable.

TaqMan MGB probes were a later refinement of the TaqMan probe technology. TaqMan MGB probes possess a minor-groove binding (MGB) molecule on the 3′ end. Where the probe binds to the target, a short minor groove is formed in the DNA, allowing the MGB molecule to bind and increase the melting temperature, thus strengthening probe binding. Consequently, TaqMan MGB probes can be much shorter than PCR primers. Because of the MGB moiety, these probes can be shorter than other TaqMan probes and still have a high melting temperature. This enables TaqMan MGB probes to bind to the target more specifically than primers at higher temperatures. With the shorter probe size, a 1-base mismatch has a much greater impact on TaqMan MGB probe binding. And because of this higher level of specificity, TaqMan MGB probes are recommended for most applications involving high genetic complexity.

TaqMan probe signal production

Whether an MGB or non-MGB probe is chosen, both follow the same pattern for signal production. In the early PCR cycles, only the low, quenched reporter signal is detected. This early signal, automatically subtracted to zero in the real-time PCR software, is termed "baseline." If the sample contains a target, eventually enough of the cleaved probe will accumulate to allow amplification signal to emerge from the baseline. The point at which the amplification signal becomes visible is inversely related to the initial target quantity.

SYBR Green dye

SYBR Green I dye is a fluorescent DNA-binding dye that binds to the minor groove of any dsDNA. Excitation of DNA-bound SYBR Green dye produces a much stronger fluorescent signal than unbound dye. A SYBR Green dye–based assay typically consists of two PCR primers. Under ideal conditions, a SYBR Green assay follows an amplification pattern similar to that of a TaqMan probe–based assay. In the early PCR cycles, a horizontal baseline is observed. If the target was present in the sample, sufficient accumulated PCR product will be produced at some point so that an amplification signal becomes visible.

SYBR Green assay specificity

Assay specificity testing is important for all assays, but especially for those most vulnerable to specificity problems. SYBR Green assays do not benefit from the specificity of a TaqMan probe, making them more vulnerable to specificity problems. SYBR Green dye will bind to any amplified product, target or non-target, and all such signals are summed, producing a single amplification plot. SYBR Green amplification plot shape cannot be used to assess specificity. Plots usually have the same appearance, whether the amplification consists of target, non-target, or a mixture. The fact that a SYBR Green assay produced an amplification product should not be automatically taken to mean the majority of the signal is derived from the target.

Since amplification of non-target can vary from sample to sample, at least one type of specificity assessment should be performed for every SYBR Green reaction. Most commonly, this ongoing assessment is dissociation analysis.

SYBR Green dye dissociation

SYBR Green dissociation is the gradual melting of the PCR products after PCR when using SYBR Green–based detection. Dissociation is an attractive choice for specificity assessment because it does not add cost to the experiment and can be done right in the PCR reaction vessel. However, dissociation does add more time to the thermal protocol, requires additional analysis time, is somewhat subjective, and has limited resolution.

The concept of SYBR Green dissociation is that if the target is one defined genetic sequence, it should have one specific T_m, which is used to help identify the target in samples. Some non-target products will have T_m values significantly different from that of the target, allowing detection of those non-target amplifications.

The dissociation protocol is added after the final PCR cycle. Following the melting process, the real-time PCR software will plot the data as the negative first derivative, which transforms the melting profile into a peak.

Accurate identification of the target peak depends on amplification of pure target. Many samples such as cellular RNA and gDNA exhibit high genetic complexity, creating opportunities for non-target amplification that may suppress the amplification of the target or, in some cases, alter the shape of the melt peak. By starting with pure target, the researcher will be able to associate a peak T_m and shape with a particular target after amplification. Only one peak should be observed. The presumptive target peak should be narrow, symmetrical, and devoid of other anomalies, such as shoulders, humps, or splits. These anomalies are strong indications that multiple products of similar T_m values were produced, casting strong doubts about the specificity of those reactions. Wells with dissociation anomalies should be omitted from further analysis.

SYBR Green dissociation is low resolution and may not differentiate between target and non-target with similar T_m values (e.g., homologs). Therefore, one narrow, symmetric peak should not be assumed to be the target, nor even a single product, without additional supporting information.

Dissociation data should be evaluated for each well where amplification was observed. If the sample contains a peak that does not correspond to the pure target peak, the conclusion is that target was not detected in that reaction. If the sample contains a peak that appears to match the T_m and shape of the pure target peak, the target may have been amplified in that reaction. Dissociation data in isolation cannot be taken as definitive, but when combined with other information, such as data from target-negative samples, sequencing, or gels, can provide more confidence in specificity.

Real-time PCR instrumentation

Many different models of real-time PCR instruments are available. Each model must have an excitation source, which excites the fluorescent dyes, and a detector to detect the fluorescent emissions. In addition, each model must have a thermal cycler. The thermal block may be either fixed, as in the Applied Biosystems™ QuantStudio™ 3, QuantStudio™ 5, or StepOnePlus™ system or user interchangeable, as in the ViiA 7 system, the Applied Biosystems™ QuantStudio 6 and 7 Flex systems, and the Applied Biosystems™ QuantStudio 12K Flex system. Blocks are available to accept a variety of PCR reaction vessels: 48-well plates, 96-well plates, 384-well plates, 384-microwell cards, 3,072–through-hole plates, etc. All real-time PCR instruments also come with software for data collection and analysis.

Dye differentiation

Most real-time PCR reactions contain multiple dyes, including one or more reporter dyes, in some cases a quencher dye, and, very often, a passive reference dye. Multiple dyes in the same well can be measured independently, either through optimized combinations of excitation and emission filters or through a process called multicomponenting.

Multicomponenting is a mathematical method to measure dye intensity for each dye in the reaction. Multicomponenting offers the benefits of easy correction for dye designation errors, refreshing of optical performance to factory standard without hardware adjustment, and providing a source of troubleshooting information.

1.6 Melting curve analysis

Melting curve analysis and detection systems

Melting curve analysis can only be performed with real-time PCR detection technologies in which the fluorophore remains associated with the amplicon. Amplifications that have used SYBR Green I or SYBR GreenER™ dye can be subjected to melting curve analysis. Dual-labeled probe detection systems such as TaqMan probes are not compatible because they produce an irreversible change in signal by cleaving and releasing the fluorophore into solution during the PCR; however, the increased specificity of this method makes this less of a concern.

The level of fluorescence of both SYBR Green I and SYBR GreenER dyes significantly increases upon binding to dsDNA. By monitoring the dsDNA as it melts, a decrease in fluorescence will be seen as soon as the DNA becomes single-stranded and the dye dissociates from the DNA.

Importance of melting curve analysis

The specificity of a real-time PCR assay is determined by the primers and reaction conditions used. However, there is always the possibility that even well-designed primers may form primer-dimers or amplify a nonspecific product (Figure 12). There is also the possibility when performing qRT-PCR that the RNA sample contains gDNA, which may also be amplified. The specificity of the real-time PCR reaction can be confirmed using melting curve analysis. When melting curve analysis is not possible, additional care must be used to establish that differences observed in C_t values between reactions are valid and not due to the presence of nonspecific products.

Melting curve analysis and primer-dimers

Primer-dimers occur when two PCR primers (either same-sense primers or sense and antisense primers) bind to each other instead of the target. Melting curve analysis can identify the presence of primer-dimers because they exhibit a lower melting temperature than the amplicon. The presence of primer-dimers is not desirable in samples that contain template, as it decreases PCR efficiency and obscures analysis. The formation of primer-dimers most often occurs in no-template controls, where there is an abundance of primer and no template. The presence of primer-dimers in the no-template control should serve as an alert to the user that they may also be present in reactions that include template. If there are primer-dimers in the no-template control, the primers should be redesigned. Melting curve analysis of no-template controls can discriminate between primer-dimers and spurious amplification due to contaminating nucleic acids in the reagent components.

How to perform melting curve analysis

To perform melting curve analysis, the real-time PCR instrument can be programmed to include a melting profile immediately following the thermal cycling protocol. After amplification is complete, the instrument will reheat the amplified products to give complete melting curve data (Figure 13). Most real-time PCR instrument platforms now incorporate this feature into their analysis packages.

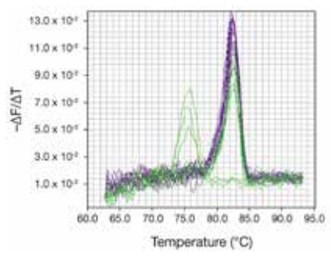


Figure 12. Melting curve analysis can detect the presence of nonspecific products, such as primer-dimers, as shown by the additional peaks to the left of the peak for the amplified product.

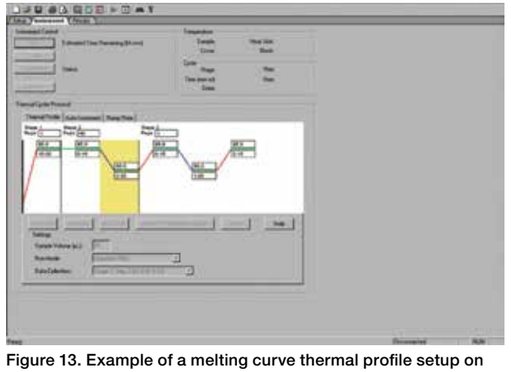


Figure 13. Example of a melting curve thermal profile setup on an Applied Biosystems instrument (rapid heating to 94°C to denature the DNA, followed by cooling to 60°C).

1.7 Passive reference dyes

Passive reference dyes are frequently used in real-time PCR to normalize the fluorescent signal of reporter dyes and correct for fluctuations in fluorescence that are not PCR-based. Normalization is necessary to correct for fluctuations from well to well caused by changes in reaction concentration or volume and to correct for variations in instrument scanning. Most real-time PCR instruments use ROX™ dye as the passive reference dye, because ROX dye does not affect real-time PCR reactions and has a fluorescent signal that can be distinguished from that of many reporter or quencher dyes used. An exception is the Bio-Rad™ iCycler iQ™ instrument system, which uses fluorescein as the reference dye.

Passive reference dye

A passive reference dye such as ROX dye is used to normalize the fluorescent reporter signal in real-time PCR on compatible instruments, such as Applied Biosystems instruments. The use of a passive reference dye is an effective tool for the normalization of fluorescent reporter signal without modifying the instrument's default analysis parameters. TaqMan™ real-time PCR master mixes contain a passive reference dye that serves as an internal control to:

- Normalize for non–PCR-related fluctuations in fluorescence (e.g., caused by variation in pipetting)
- Normalize for fluctuations in fluorescence resulting from machine "noise"
- Compensate for variations in instrument excitation and detection
- Provide a stable baseline for multiplex real-time PCR and qRT-PCR

Fluorescein reference dye

Bio-Rad™ iCycler™ instruments require the collection of "well factors" before each run to compensate for any instrument or pipetting nonuniformity. Well factors for experiments using SYBR Green I or SYBR GreenER dye are calculated using an additional fluorophore, fluorescein.

Well factors are collected using either a separate plate containing fluorescein dye in each well (for external well factors) or the experimental plate with fluorescein spiked into the real-time PCR master mix (for dynamic well factors). The method must be selected when starting each run on the iCycler instrument. The iCycler iQ™5 and MyiQ™ systems allow saving the data from an external well factor reading as a separate file, which can then be referenced for future readings.

1.8 Contamination prevention

As with traditional PCR, real-time PCR can be affected by nucleic acid contamination, leading to false positive results. Some of the possible sources of contamination are:

- Cross-contamination between samples
- Contamination from laboratory equipment
- Carryover contamination of amplification products and primers from previous PCR runs. This is considered to be the major source of false-positive PCR results

Uracil N-glycosylase (UNG)

UNG is used to reduce or prevent DNA carryover contamination between PCR reactions by preventing the amplification of DNA from previous reactions. The use of UNG in PCR reduces false positives, in turn increasing the efficiency of the real-time PCR reaction and the reliability of data.

How UNG carryover prevention works

UNG for carryover prevention begins with the substitution of dUTP for dTTP in real-time PCR master mixes. Subsequent real-time PCR reaction mixes are then treated with UNG, which degrades any contaminating uracil-containing PCR products, leaving the natural (thymine-containing) target DNA template unaffected.

With standard UNG, a short incubation at 50°C is performed prior to the PCR thermal cycling so that the enzyme can cleave the uracil residues in any contaminating DNA. The removal of the uracil bases causes fragmentation of the DNA, preventing its use as a template in PCR. The UNG is then inactivated in the ramp-up to 95°C in PCR. A heat-labile form of the enzyme is also available, which is inactivated at 50°C, so that it can be used in one-step qRT-PCR reaction mixes.

1.9 Multiplex real-time PCR

Introduction to multiplexing

PCR multiplexing is the amplification and specific detection of two or more genetic sequences in the same reaction. To be successful, PCR multiplexing must be able to produce sufficient amplified product for the detection of all of the intended sequences. Real-time PCR multiplexing may be used to produce quantitative or qualitative results. For quantitative PCR multiplexing, all of the intended sequences must produce sufficient exponential phase signal. For qualitative results, if amplified products are sufficient, an endpoint detection method such as gel electrophoresis can be used.

The suffix “plex” is used in multiple terms. A singleplex assay is designed to amplify a single genetic sequence. A duplex is a combination of two assays designed to amplify two genetic sequences. The most common type of multiplex is a duplex, in which the assay for the target gene is conducted in the same well as that for the control or normalizer gene, but higher-order multiplexes are also possible.

Some commercial real-time PCR kits are designed and validated as a multiplex. For example, the Applied Biosystems™ MicroSEQ™ *E. coli* O157:H7 Kit multiplexes the *E. coli* target assay with an internal positive control assay. For research applications, the scientist usually chooses which assays to multiplex and is responsible for multiplex validation. When considering whether to create a multiplex assay, it is important to weigh the benefits of multiplexing against the degree of effort needed for validation.

Multiplexing benefits

Three benefits of multiplexing—increased throughput (more samples potentially assayed per plate), reduced sample usage, and reduced reagent usage—are dependent on the number of targets in the experiment. For example, if a quantitative experiment consists of only one target assay, running the target assay as a duplex with the normalizer assay, such as an endogenous control assay, will increase throughput, reduce sample required, and reduce reagent usage by half. If a quantitative experiment consists of two target assays, it may be possible to combine two target assays and the normalizer assay in a triplex reaction. In that case, the throughput increase, sample reduction, and reagent reduction will be even greater.

If the target assay is multiplexed with the normalizer assay, another benefit of multiplexing is minimizing pipette precision errors. Target and normalizer data from the same well are derived from a single sample addition, so any pipette precision error should affect both the target and normalizer results equally. In order to gain this precision benefit, target data must be normalized by the normalizer data from the same well before calculating technical replicate precision. Comparing multiplex data analyzed in a singleplex manner (without well-based normalization) to an analysis done in a multiplex manner demonstrates that the multiplex precision benefit can be substantial, depending on the singleplex error. For example, for samples with minimal singleplex precision error, the multiplex precision benefit will be minimal as well.

The precision benefit of multiplexing is especially valuable for quantitative experiments requiring a higher degree of precision. For example, in copy number variation experiments, discriminating 1 copy from 2 copies of the gene is a 2-fold difference, which requires good precision. However, discriminating 2 copies from 3 copies is only a 1.5-fold difference, which requires even better precision. Multiplexing is one recommended method to help achieve the necessary degree of precision for this type of experiment.

Instrumentation for multiplexing

Multiplex assays usually involve multiple dyes in the same well. The real-time PCR instrument must be capable of measuring those different dye signals in the same well with accuracy. These measurements must remain specific for each dye, even when one dye signal is significantly higher than another. Proper instrument calibration is necessary to accurately measure each dye contribution within a multiplex assay.

Chemistry recommendations for multiplexing

The best fluorescent chemistries for real-time PCR multiplexing are those that can assign different dyes to detect each genetic sequence in the multiplex. The vast majority of multiplexing is performed with multi-dye, high-specificity chemistries, such as TaqMan probe–based assays.

For multiplex assays involving RNA, two-step RT-PCR is generally recommended over one-step RT-PCR. One-step RT-PCR requires the same primer concentration for reverse transcription and PCR, reducing flexibility in primer concentrations optimal for multiplexing. In two-step RT-PCR, the PCR primer concentration may be optimized for multiplexing without having any adverse affect on reverse transcription.

Dye choices for multiplexing

Assuming a multi-dye real-time PCR fluorescence chemistry is being used, each genetic sequence being detected in the multiplex assay will require a different reporter dye. The reporter dyes chosen must be sufficiently excited and accurately detected when together in the same well by the real-time PCR instrument. The instrument manufacturer should be able to offer dye recommendations. Note that Applied Biosystems real-time PCR master mixes contain a red passive reference dye. Whereas blue-only excitation instruments can excite this ROX dye–based reference sufficiently to act as a passive reference dye, blue excitation is generally not sufficient for red dyes to act as a reporter.

Reporter dyes do not have to be assigned based on the type of target gene or gene product, but following a pattern in assigning dyes can simplify the creation of a multiplex assay. For example, FAM dye is the most common reporter dye used in TaqMan probes, so following the pattern of assigning FAM dye as the reporter for the target assay, VIC™ dye would be the reporter for the normalizer assay. Using this pattern, multiple duplex assays may be created by pairing a different FAM dye–labeled target assay with the same VIC dye for the normalizer assay. In a triplex assay, a third dye, such as ABY™ dye or JUN™ dye, may be combined with FAM dye and VIC dye. Note that if ABY dye is being used, TAMRA dye should not be present in the same well, and if JUN dye is used, MUSTANG PURPLE™ dye should be used instead of ROX dye as a passive reference.

Multiplex PCR saturation

Multiplex PCR saturation is an undesirable phenomenon that may occur in a multiplex assay when the amplification of the more abundant gene saturates the thermostable DNA polymerase, suppressing the amplification of the less abundant gene. The remedy for saturation is a reduction of the PCR primer concentration for the more abundant target, termed “primer limitation.” The primer-limited concentration should be sufficient to enable exponential amplification, but sufficiently low that the primer is exhausted before the PCR product accumulates to a level that starves amplification of the less abundant target.

When planning a multiplex assay, the researcher should identify which gene or genes in the multiplex have the potential to cause saturation, which is based on the absolute DNA or cDNA abundance of each gene or gene product in the PCR reaction. In this regard, the three most common duplex scenarios are listed below.

Duplex scenario 1

In this most common scenario, the more abundant gene or gene product is the same in all samples. Only the assay for the more abundant target requires primer limitation. For example, the normalizer might be 18S ribosomal RNA (rRNA), which is 20% of eukaryotic total RNA. The 18S rRNA cDNA would be more abundant than any cDNA derived from mRNA in every sample. Therefore, only the 18S assay would require primer limitation.

Duplex scenario 2

In this scenario, the two genes have approximately equal abundance in all samples. Generally, a C_t difference between the two genes of 3 or higher, assuming the same threshold, would not qualify for equal abundance. gDNA applications, such as copy number variation, are most likely to fall into this scenario. Primer limitation is not necessary for scenario 2, because the two genes are progressing through the exponential amplification phase at approximately the same time.

Duplex scenario 3

In this scenario, either the gene or gene product has the potential to be significantly more abundant than the other. In this case, both assays should be primer limited.

Multiplex primer interactions

Another potential threat to multiplex assay performance is unexpected primer interactions between primers from different assays. The risk of primer interaction grows with the number of assays in the reaction, because the number of unique primer pairs increases dramatically with the number of assays in the multiplex. For example, in a duplex assay with 4 PCR primers there are 6 unique primer pairs possible, and in a triplex assay with 6 PCR primers there are 15 unique pairs possible. In singleplex, each assay may perform well, but in a multiplex reaction the primer interactions can create competitive products, suppressing amplification. The chances of primer interactions grow when the assays being multiplexed have homology. If primer interaction does occur based on the observation of significantly different C_t values in the singleplex vs. multiplex reaction, the remedy is to use a different assay in the multiplex reaction.

1.10 Internal controls and reference genes

Real-time PCR has become a method of choice for gene expression analysis. To achieve accurate and reproducible expression profiling of selected genes using real-time PCR, it is critical to use reliable internal control gene products for the normalization of expression levels between experiments—typically expression products from housekeeping genes are used. The target chosen to be the internal standard (or endogenous control) should be expressed at roughly the same level as the experimental gene product. By using an endogenous control as an active reference, quantification of an mRNA target can be normalized for differences in the amount of total RNA added to each reaction. Regardless of the gene that is chosen to act as the endogenous control, that gene must be tested to ensure that there is consistent expression of the control gene under all desired experimental conditions.

Relative gene expression analysis using housekeeping genes

Relative gene expression comparisons work best when the expression level of the chosen housekeeping gene remains constant. The choice of the housekeeping reference gene is reviewed in *BioTechniques* 29:332 (2000) and *J Mol Endocrinol* 25:169 (2000). Ideally, the expression level of the chosen housekeeping gene should be validated for each target cell or tissue type to confirm that it remains constant at all points of the experiment. For example, GAPDH expression has been shown to be upregulated in proliferating cells, and 18S rRNA may not always represent the overall cellular mRNA population.

1.11 Real-time PCR instrument calibration

Timely, accurate calibration is critical for the proper performance of any real-time PCR instrument. It preserves data integrity and consistency over time. Real-time PCR instruments should be calibrated as part of a regular maintenance regimen and prior to using new dyes for the first time, following the manufacturer's instructions.

Excitation/emission difference corrections

The optical elements in real-time PCR instruments can be divided into two main categories: the excitation source, such as halogen lamps or LEDs, and the emission detector, such as a CCD camera or photodiode. While manufacturers can achieve excellent uniformity for excitation strength and emission sensitivity across the wells of the block, there will always be some variation. This variation may increase with age and usage of the instrument. Uncorrected excitation/emission differences across the plate can cause shifts in C_t values. However, if a passive reference dye is present in the reaction, those differences will affect the reporter and passive reference signals to the same degree, so that normalization of the reporter to the passive reference corrects the difference.

Universal optical fluctuations

In traditional plastic PCR plates and tubes, the liquid reagents are at the bottom of the well, an air space is above the liquid, and a plastic seal is over the well. With this configuration, a number of temperature-related phenomena occur.

During cycling, temperatures reach 95°C. At that high temperature, water is volatilized into the air space in the well. This water vapor or steam will condense on the cooler walls of the tube, forming water droplets that return to the reagents at the bottom. This entire process, called "refluxing," is continuous during PCR.

Second, at high temperature, air dissolved within the liquid reagents will become less soluble, creating small air bubbles.

Third, the pressure of the steam will exert force on the plastic seal, causing it to change shape slightly during PCR.

All of these temperature-related phenomena are in the excitation and emission light path and can cause fluctuations in fluorescent signal. The degree of these fluctuations can vary, depending on factors such as how much air was dissolved in the reagents and how well the plate was sealed. Generally, universal fluctuations do not produce obvious distortions in the reporter signal, but they do affect the precision of replicates. If present, a passive reference dye is in the same light path as the reporter, so normalization of reporter to passive reference signals corrects for these fluctuations.

Precision improvement

The correction effect of passive reference normalization will improve the precision of real-time PCR data. The degree of improvement will vary, depending on a number of factors, such as how the reagents and plate were prepared.

Atypical optical fluctuations

Atypical optical fluctuations are thermal-related anomalies that are not universal across all reactions in the run and produce an obvious distortion in the reporter signal. One example of an atypical optical fluctuation is a significant configuration change in the plate seal, which may be termed "optical warping." Optical warping occurs when a well is inadequately sealed and then, during PCR, the heat and pressure of the heated lid causes the seal to seat properly. A second example is large bubbles that burst during PCR.

Distortions in the amplification plot are likely to cause baseline problems and may even affect C_t values. Normalization to a passive reference dye provides excellent correction for optical warping, so the resulting corrected amplification plot may appear completely anomaly free. Normalization does not fully correct for a large bubble bursting, but it can help minimize the data distortion caused.

Experimental design

2

Contents

2.1 Introduction

Successful real-time PCR assay design and development are the foundation for accurate data. Up-front planning will assist in managing any experimental variability observed during this process.

Before embarking on experimental design, clearly understand the goal of the assay; specifically, what biological questions need to be answered. For example, an experiment designed to determine the relative expression level of a gene in a particular disease state will be quite different from one designed to determine viral copy number from that same disease state. After determining the experimental goal, identify the appropriate real-time PCR controls and opportunities for optimization.

This section describes the stages of real-time PCR assay design and implementation. It will identify sources of variability and the role they play in data accuracy, and provide guidelines for optimization in the following areas:

- Target amplicon and primer design
- Nucleic acid purification
- Reverse transcription
- Controls and normalization
- Standard curve evaluation of efficiency, sensitivity, and reproducibility

2.2 Real-time PCR assay types

Gene expression profiling is a common use of real-time PCR that assesses the relative abundance of transcripts to determine gene expression patterns between samples. RNA quality, reverse transcription efficiency, real-time PCR efficiency, quantification strategy, and the choice of a normalizer gene play particularly important roles in gene expression experiments.

Viral titer determination assays can be complex to design. Often, researchers want to quantify viral copy numbers in samples. This is often accomplished by comparison to a standard curve generated using known genome equivalents or nucleic acid harvested from a titered virus control. Success is dependent on the accuracy of the material used to generate the standard curve. Depending on the nature of the target—an RNA or DNA virus—reverse transcription and real-time PCR efficiency also play significant roles. Assay design will also be influenced by whether the assay is intended to count functional viral particles or the total number of particles.

Copy number variation analysis analyzes the genome for duplications or deletions. The assay design, and most specifically standard curve generation, will be dictated by whether relative or absolute quantification is desired. Assay design focuses on real-time PCR efficiency and the accuracy necessary to discriminate single-copy deviations.

Allelic discrimination assays can detect variation down to the single-nucleotide level. Unlike the methods described above, endpoint fluorescence is measured to determine SNP genotypes. Primer and probe design play particularly important roles to ensure a low incidence of allele-specific cross-reactivity.

2.3 Amplicon and primer design considerations

Target amplicon size, GC content, location, and specificity

As will be discussed in more detail later in this guide, reaction efficiency is paramount to the accuracy of real-time PCR data. In a perfect scenario, each target copy in a PCR reaction will be copied at each cycle, doubling the number of full-length target molecules: this corresponds to 100% amplification efficiency. Variations in efficiency will be amplified as thermal cycling progresses. Thus, any deviation from 100% efficiency can result in potentially erroneous data.

One way to minimize efficiency bias is to amplify relatively short targets. Amplifying a 100 bp region is much more likely to result in complete synthesis in a given cycle than, say, amplifying a 1,200 bp target. For this reason, real-time PCR target lengths are generally 60–200 bp. In addition, shorter amplicons are less affected by variations in template integrity. If nucleic acid samples are slightly degraded and the target sequence is long, upstream and downstream primers will be less likely to find their complementary sequence in the same DNA fragment.

Amplicon GC content and secondary structure can be another cause of data inaccuracy. Less-than-perfect target doubling at each cycle is more likely to occur if secondary structure obstructs the path of the DNA polymerase. Ideally, primers should be designed to anneal with, and to amplify, a region of medium (50%) GC content with no significant GC stretches. For amplifying cDNA, it is best to locate amplicons near the 3′ ends of transcripts. If RNA secondary structure prohibits full-length cDNA synthesis in a percentage of the transcripts, these amplicons are less likely to be impacted (Figure 14).

Target specificity is another important factor in data accuracy. When designing real-time PCR primers, check primers to be sure that their binding sites are unique in the genome. This reduces the possibility that the primers could amplify similar sequences elsewhere in the sample genome. Primer design software programs automate the process of screening target sequences against the originating genome and masking homologous areas, thus eliminating primer designs in these locations.

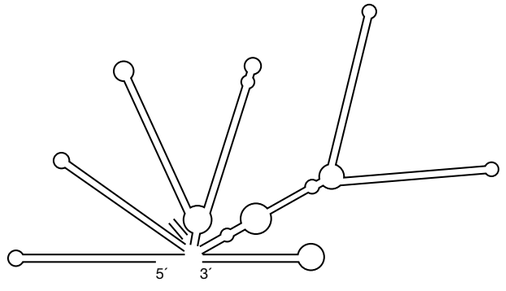


Figure 14. An RNA molecule with a high degree of secondary structure.

gDNA, pseudogenes, and allelic variants

gDNA carryover in an RNA sample may be a concern when measuring gene expression levels. The gDNA may be co-amplified with the target transcripts of interest, resulting in invalid data. gDNA contamination is detected by setting up control reactions that do not contain reverse transcriptase (no-RT controls); if the C_t for the no-RT control is higher than the C_t generated by the most dilute target, it indicates that gDNA is not contributing to signal generation. However, gDNA can compromise the efficiency of the reaction due to competition for reaction components such as dNTPs and primers.

The best method for avoiding gDNA interference in real-time PCR is thoughtful primer (or primer/probe) design, which takes advantage of the introns present in gDNA that are absent in mRNA. Whenever possible, Applied Biosystems™ TaqMan™ Gene Expression Assays are designed so that the TaqMan probe spans an exon-exon boundary. Primer sets for SYBR Green dye–based detection should be designed to anneal in adjacent exons or with one of the primers spanning an exon/exon junction. When upstream and downstream PCR primers anneal within the same exon, they can amplify target from both DNA and RNA. Conversely, when primers anneal in adjacent exons, only cDNA will be amplified in most cases, because the amplicon from gDNA would include intron sequence, resulting in an amplicon that is too long to amplify efficiently in the conditions used for real-time PCR.

Pseudogenes, or silent genes, are other transcript variants to consider when designing primers. These are derivatives of existing genes that have become nonfunctional due to mutations and/or rearrangements in the promoter or gene itself. Primer design software programs can perform BLAST™ searches to avoid pseudogenes and their mRNA products.

Allelic variants are two or more unique forms of a gene that occupy the same chromosomal locus. Transcripts originating from these variants can vary by one or more mutations. Allelic variants should be considered when

designing primers, depending on whether one or more variants are being studied. In addition, GC-content differences between variants may alter amplification efficiencies and generate separate peaks on a melt curve, which can be incorrectly diagnosed as off-target amplification. Alternately spliced variants should also be considered when designing primers.

Specificity, dimerization, and self-folding in primers and probes

Primer-dimers are most often caused by an interaction between forward and reverse primers, but can also be the result of forward-forward or reverse-reverse primer annealing, or a single primer folding upon itself. Primer-dimers are of greater concern in more complex reactions such as multiplex real-time PCR. If the dimerization occurs in a staggered manner, as often is the case, some extension can occur, resulting in products that approach the size of the intended amplicon and become more abundant as cycling progresses. Typically, the lower the amount of target at the start of the PCR, the more likely primer-dimer formation will be. The positive side of this potential problem is that the interaction of primer-dimers is usually less favorable than the intended primer-template interaction, and there are many ways to minimize or eliminate this phenomenon.

The main concern with primer-dimers is that they may cause false-positive results. This is of particular concern with reactions that use DNA-binding dyes such as SYBR Green I dye. Another problem is that the resulting competition for reaction components can contribute to a reaction efficiency outside the desirable range of 90–110%. The last major concern, also related to efficiency, is that the dynamic range of the reaction may shrink, impacting reaction sensitivity. Even if signal is not generated from the primer-dimers themselves (as is the case with TaqMan Assays), efficiency and dynamic range may still be affected.

Several free software programs are available to analyze real-time PCR primer designs and determine if they will be prone to dimerize or fold upon themselves. The AutoDimer software program (authored by P.M. Vallone, National Institute of Standards and Technology, USA) is a bioinformatics tool that can analyze a full list of primers at the same time (Figure 15). This is especially helpful with multiplexing applications. However, while bioinformatics analysis of primer sequences can greatly minimize the risk of dimer formation, it is still necessary to monitor dimerization experimentally.

Figure 15. A screen capture from AutoDimer software. This software is used to analyze primer sequences and report areas of potential secondary structure within primers (which could lead to individual primers folding on themselves) or stretches of sequence that would allow primers to anneal to each other.

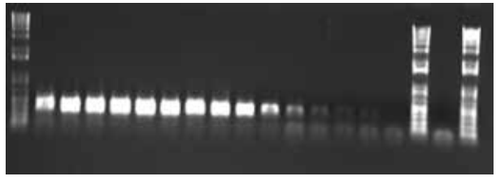


Figure 16. Agarose gel analysis to investigate primer-dimer formation. Prior to the thermal cycling reaction, the nucleic acid sample was serially diluted and added to the components of a PCR mix, and the same volume from each mixture was loaded on an agarose gel. Primer-dimers appear as diffuse bands at the bottom of the gel.

The traditional method of screening for primer-dimers is gel electrophoresis. Primer-dimers appear as diffuse, smudgy bands near the bottom of the gel (Figure 16). One concern with gel validation is that it is not very sensitive and therefore may be inconclusive. However, gel analysis is useful for validating data obtained from a melting/dissociation curve, which is considered the best method for detecting primer-dimers.

Melting or dissociation curves should be generated following any real-time PCR run that uses DNA-binding dyes for detection. In brief, the instrument ramps from low temperature, in which DNA is double-stranded and fluorescence is high, to high temperature, which denatures DNA and results in lower fluorescence. A sharp decrease in fluorescence will be observed at the T_m for each product generated during the PCR. The melting curve peak obtained for the no-template control can be compared to the peak obtained from the target to determine whether primer-dimers are present in the reaction.

Ideally, a single distinct peak should be observed for each reaction containing template, and no peaks should be present in the no-template controls. Smaller, broader peaks at a lower melting temperature than that of the desired amplicon and also appearing in the no-template control reactions are quite often dimers. Again, gel runs of product can often validate the size of the product corresponding to the melting peak.

There are situations in which primer-dimers are present, but they may not affect the overall accuracy of the real-time PCR assay. A common observation is that primer-dimers are present in the no-template control but do not appear in reactions containing template DNA. This is not surprising because in the absence of template, primers are much more likely to interact with each other. When template is present, primer-dimer formation is not favored. As long as the peak seen in the no-template control is absent in the plus-template dissociation curve, primer-dimers are not an issue.

Primer-dimers are part of a broad category of nonspecific PCR products that includes amplicons created when a primer anneals to an unexpected location with an imperfect match. Amplification of nonspecific products is of concern because they can contribute to fluorescence, which in turn artificially shifts the C_t of the reaction. They can influence reaction efficiency through competition for reaction components, resulting in a decreased dynamic range and decreased data accuracy. Nonspecific products are an even greater concern in absolute quantification assays, in which precise copy numbers are reported.

Standard gel electrophoresis is generally the first step in any analysis of real-time PCR specificity. While it can help to identify products that differ in size from a target amplicon, a band may still mask similar-sized amplicons and have limited sensitivity. Due to its accuracy and sensitivity, melting curve analysis provides the most confidence in confirming gel electrophoretic assessment of primer specificity.

While nonspecific amplification should always be eliminated when possible, there are some cases in which the presence of these secondary products is not a major concern. For example, if alternate isoforms or multiple alleles that differ in GC content are knowingly targeted, multiple products are expected.

Primer design considerations

The following recommendations are offered for designing primers for real-time PCR: Applied Biosystems™ Primer Express™ Software, Invitrogen™ OligoPerfect™ Designer web-based tool, and Invitrogen™ Vector NTI™ Software. Note that primer design software programs, such as our web-based OligoPerfect Designer and Vector NTI Software, are seamlessly connected to our online ordering system, so there's no need to cut and paste sequences. These programs can automatically design primers for specific genes or target sequences using algorithms that incorporate the following guidelines and can also perform genome-wide BLAST searches for known sequence homologies.

- In general, design primers that are 18–28 nucleotides in length
- Avoid stretches of repeated nucleotides
- Aim for 50% GC content, which helps to prevent mismatch stabilization
- Choose primers that have compatible T_m values (within 1°C of each other)
- Avoid sequence complementarity between all primers employed in an assay and within each primer

2.4 Nucleic acid purification and quantitation

Real-time PCR nucleic acid purification methods

Prior to purifying nucleic acid, the source material (cells or tissue) and potential technique limitations must be considered. DNA and RNA isolation techniques vary in ease of use, need for organic solvents, and resulting nucleic acid purity with regards to carryover of DNA (in the case of RNA isolation), protein, and organic solvents.

This section will primarily discuss RNA isolation, though most of the same guidelines also hold true for DNA isolation.

One-step reagent-based organic extraction is a very effective method for purifying RNA from a wide variety of cell and tissue types. Many protocols use a phenol and guanidine isothiocyanate mixture to disrupt cells and dissolve cell components while maintaining the integrity of the nucleic acids by protecting them from RNases. Guanidine isothiocyanate is a chaotropic salt that protects RNA from endogenous RNases (*Biochemistry* 18:5294 (1979)). Typically, chloroform is then added and the mixture is separated into aqueous and organic phases by centrifugation. RNA remains exclusively in the aqueous phase in the presence of guanidine isothiocyanate, while DNA and protein are driven into the organic phase and interphase. The RNA is then recovered from the aqueous phase by precipitation with isopropyl alcohol.

This process is relatively fast and can yield high levels of RNA, but requires the use of toxic chemicals and may result in higher DNA carryover compared to other techniques. Residual guanidine, phenol, or alcohol can also dramatically reduce cDNA synthesis efficiency.

With most methods based on silica beads or filters, samples are lysed and homogenized in the presence of guanidine isothiocyanate. After homogenization, ethanol is added to the sample, RNA is bound to silica-based beads or filters, and impurities are effectively removed by washing (*Proc Natl Acad Sci USA* 76:615 (1979)). The purified total RNA is eluted in water.

This method is even less time-consuming than organic extraction and does not require phenol. The RNA yield may not be quite as high, but the purity with regards to protein, lipids, polysaccharides, DNA, and purification reagents is generally better. Guanidine and ethanol carryover due to incomplete washing can still occur, with the same deleterious effects on cDNA synthesis efficiency.

Lastly, methods combining organic lysis with silica columns can offer the benefits of good sample lysis with the ease, speed, and purity of silica-binding methods.

Assessing RNA quality

In assessing RNA quality and quantity, there are a few key points to focus on. Ensure that the ratio of UV absorbance at 260 nm to the absorbance at 280 nm (A_{260}/A_{280} ratio) is between 1.8 and 2.0. A ratio below 1.8 can indicate protein contamination, which can lower reaction efficiency. The A_{260}/A_{230} ratio is helpful in evaluating the carryover of components containing phenol rings such as the chaotropic salt guanidine isothiocyanate and phenol itself, which are inhibitory to enzymatic reactions. Assess RNA integrity on a denaturing gel or on an instrument such as the Agilent™ Bioanalyzer™ system (Figure 17).

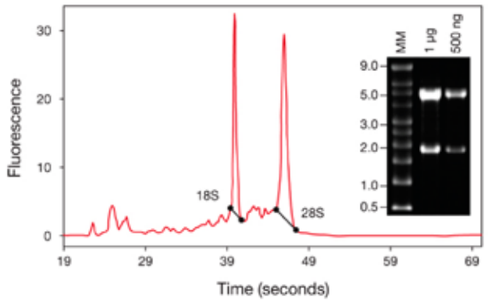


Figure 17. Agilent Bioanalyzer system trace and gel image displaying RNA integrity. Intact mammalian total RNA shows two bands or peaks representing the 18S and 28S rRNA species. In general, the 28S rRNA is twice as bright (or has twice the area under the peak in the Bioanalyzer system trace) as the 18S rRNA. MM: Invitrogen™ Millenium™ RNA Markers.

The Bioanalyzer system takes RNA quality determination one step further with the assignment of an RNA integrity number (RIN) value. The RIN value is calculated from the overall trace, including degraded products, which in general is better than assessing the rRNA peaks alone.

Researchers are then able to compare RIN values for RNA from different tissue types to assess standardization of quality and maintenance of consistency.

Quantitation accuracy

For quantitation of RNA, fluorescent dyes such as Invitrogen™ RiboGreen™ and PicoGreen™ dyes are superior to UV absorbance measurements because they are designed to have higher sensitivity, higher accuracy, and high-throughput capability. UV absorbance measurements cannot distinguish between nucleic acids and free nucleotides. In fact, free nucleotides absorb more at 260 nm than do nucleic acids. Similarly, UV absorbance measurements cannot distinguish between RNA and DNA in the same sample. In addition, contaminants commonly present in samples of purified nucleic acid contribute to UV absorbance readings. Finally, most UV absorbance readers consume a considerable amount of the sample for the measurement itself. With the wide variety of fluorescent dyes available, it is possible to find reagents that overcome all of these limitations: dyes that can distinguish nucleic acids from free nucleotides, dyes that can distinguish DNA from RNA in the same sample, and dyes that are insensitive to common sample contaminants. The Invitrogen™ Qubit™ quantitation platform uses Quant-iT™ fluorescence technology, with advanced fluorophores that become fluorescent upon binding to DNA, RNA, or protein. This specificity enables more accurate results than with UV absorbance readings, because Quant-iT™ assay kits only report the concentration of the molecule of interest (not contaminants). And, in general, quantitation methods using fluorescent dyes are very sensitive and only require small amounts of sample.

gDNA carryover in expression studies

A previous discussion described how primer design is the first step toward eliminating DNA amplification in real-time RT-PCR. DNase treatment of the sample at the RNA isolation stage is a method by which DNA can be controlled at the source. An option in addition to traditional DNase I is super-active Invitrogen™ TURBO™ DNase, which is catalytically superior to wild type DNase I. It can remove even trace quantities of DNA, which can plague RT-PCR. DNase treatment can occur either in solution or on column, depending on the isolation method. On-column DNase treatments are common with silica matrix extraction, and, unlike in-solution treatments, they do not need to be heat-inactivated in the presence of EDTA, because salt washes remove the enzyme. The drawback is that on-column reactions require much more enzyme.

In-solution DNase reactions have traditionally required heat-inactivation of the DNase at 65°C. Free magnesium, required for the reaction, can cause magnesium-dependent RNA hydrolysis at this temperature. Invitrogen™ DNA-*free*™ and TURBO DNA-*free*™ kits help circumvent these problems by using a unique DNase inactivation reagent. In addition to removing DNase from reactions, the inactivation reagent also binds and removes divalent cations from the reaction buffer. This alleviates concerns about introducing divalent cations into RT-PCR reactions, where they can affect reaction efficiency.

2.5 Reverse transcription considerations

Reverse transcriptases

Most reverse transcriptases employed in qRT-PCR are derived from avian myeloblastosis virus (AMV) or Moloney murine leukemia virus (M-MLV). Native AMV reverse transcriptase is generally more thermostable than M-MLV, but produces lower yields. However, manipulations of these native enzymes have resulted in variants with ideal properties for qRT-PCR. An ideal reverse transcriptase will exhibit the following attributes:

- **Thermostability**—as discussed earlier, secondary structure can have a major impact on the sensitivity of a reaction. Native reverse transcriptases perform ideally between 42°C and 50°C, whereas thermostable reverse transcriptases function at the higher end of (or above) this range and allow for successful reverse transcription of GC-rich regions.
- **Reduced RNase H activity**—the RNase H domain is present in common native reverse transcriptases and functions *in vivo* to cleave the RNA strand of RNA-DNA heteroduplexes for the next stage of replication. For qRT-PCR applications, RNase H activity can drastically reduce the yield of full-length cDNA, which translates to poor sensitivity. Several reverse transcriptases, most notably Invitrogen™ SuperScript™ VILO™, II and III Reverse Transcriptases, have been engineered for reduced RNase H activity.

One-step and two-step qRT-PCR

The choice between one-step and two-step qRT-PCR comes down to convenience, sensitivity, and assay design. The advantages and disadvantages of each technique must be evaluated for each experiment.

In a one-step reaction, the reverse transcriptase and thermostable DNA polymerase are both present during reverse transcription, and the reverse transcriptase is inactivated in the high-temperature DNA polymerase activation stage (the so-called hot start). Normally, the reverse transcriptase is favored by a buffer that is not optimal for the DNA polymerase. Thus, one-step buffers are a compromise solution that provide acceptable but not optimal functionality of both enzymes. This slightly lower functionality is compensated by the fact that, using this single-tube procedure, all cDNA produced is amplified in the PCR stage.

The benefits of one-step qRT-PCR include the following:

- **Contamination prevention**—the closed-tube system prevents introduction of contaminants between the reverse transcription and PCR stages
- **Convenience**—the number of pipetting steps is reduced and hands-on time is minimized
- **High-throughput sample screening**—for the reasons mentioned above
- **Sensitivity**—one-step reactions may be more sensitive than two-step reactions because all the first-strand cDNA created is available for real-time PCR amplification

The drawbacks of one-step qRT-PCR include:

- **Increased risk of primer-dimer formation**—forward and reverse gene-specific primers, present from the start in one-step reactions, have a greater tendency to dimerize under the 42–50°C reverse transcription conditions. This can be especially problematic in reactions that use DNA-binding dyes for detection
- **cDNA is not available for other real-time PCR reactions**—one-step reactions use all the cDNA from the reverse transcription step, so if the reaction fails, the sample is lost

In two-step qRT-PCR, the reverse transcription is performed in a buffer optimized for the reverse transcriptase. Once complete, approximately 10% of the cDNA is transferred into each real-time PCR reaction, also in its optimal buffer.

The benefits of two-step qRT-PCR include:

- **cDNA may be archived and used for additional real-time PCR reactions**—two-step qRT-PCR produces enough cDNA for multiple real-time PCR reactions, making it optimal for rare or limited samples
- **Sensitivity**—two-step reactions may be more sensitive than one-step reactions because the reverse transcription and real-time PCR reactions are performed in their individually optimized buffers
- **Multiple targets**—depending on the reverse transcription primers used, multiple targets can be interrogated from a single RNA sample

The drawbacks of two-step qRT-PCR include:

- **Reverse transcriptases and buffers can inhibit real-time PCR**— typically, only 10% of the cDNA synthesis reaction is used in real-time PCR, because the reverse transcriptase and associated buffer components may inhibit the DNA polymerase if not diluted properly. The specific level of inhibition will depend on the reverse transcriptase, the relative abundance of the target, and the robustness of the amplification reaction.
- **Less convenience**—two-step reactions require more handling and are less amenable to high-throughput applications
- **Contamination risk**—increased risk of contamination due to the use of separate tubes for each step

RNA priming strategies

Reverse transcription is typically the most variable step in qRT-PCR. The first-strand synthesis reaction can use gene-specific, oligo(dT), or random primers (Figure 18), and primer selection can play a large role in reverse transcription efficiency and consistency and, consequently, data accuracy.

Random primers are great for generating large pools of cDNA, and therefore can offer the highest sensitivity in real-time PCR. They are also ideal for non-polyadenylated RNA, such as bacterial RNA. Because they anneal throughout the target molecule, degraded transcripts and secondary structure do not pose as much of a problem as they do with gene-specific primers and oligo(dT) primers.

While increased yield is a benefit, data have shown that random primers can overestimate copy number. Employing a combination of random and oligo(dT) primers can sometimes increase data quality by combining the benefits of both in the same reverse transcription reaction. Random primers are used only in two-step qRT-PCR.

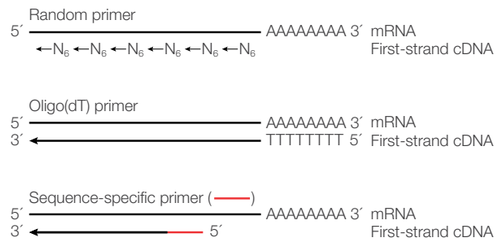


Figure 18. Graphical representation of commonly used RNA-priming strategies.

Oligo(dT) primers are a favorite choice for two-step reactions because of their specificity for mRNA and because many different targets can be analyzed from the same cDNA pool when they are used to prime reactions. However, because they always initiate reverse transcription at the 3′ end of the transcript, difficult secondary structure may lead to incomplete cDNA generation. Oligo(dT) priming of fragmented RNA, such as that isolated from formalin-fixed, paraffin-embedded samples, may also be problematic. Nonetheless, as long as the real-time PCR primers are designed near the 3′ end of the target, premature termination downstream of this location is not generally an important concern.

Multiple types of oligo(dT) primers are available. $Oligo(dT)_{20}$ is a homogeneous mixture of 20-mer thymidines, while $oligo(dT)_{12–18}$ is a mixture of 12-mer to 18-mer thymidines. Anchored oligo(dT) primers are designed to avoid poly(A) slippage by annealing at the 3′ untranslated region (UTR)/poly(A) junction. Choosing the best oligo(dT) primer may depend in part on the reaction temperature. More thermostable reverse transcriptases may perform better with longer primers, which remain more tightly annealed at elevated temperatures compared to their shorter counterparts. Oligo(dT) primers are not recommended if 18S rRNA is used for normalization.

Sequence-specific primers offer the greatest specificity and have been shown to be the most consistent of the primer options for reverse transcription. However, they do not offer the flexibility of oligo(dT) and random primers, in that they only produce a cDNA copy of the target gene product. Because of this, gene-specific primers are typically not the best choice for studies involving scarce or precious samples. One-step qRT-PCR always employs a gene-specific primer for first-strand synthesis, whereas other primer options are compatible with two-step reactions.

Each primer type presents unique theoretical benefits and drawbacks. In addition, individual targets may respond differently to one primer choice over another. Ideally, each primer option should be evaluated during the initial assay validation stage to determine which provides optimal sensitivity and accuracy.

Factors influencing reverse transcription efficiency

The reverse transcription stage of a qRT-PCR reaction is less consistent than the PCR stage. This is due to a combination of factors associated with the starting

sample, which the thermostable DNA polymerase isn't normally tested with. These factors include:

Differences in RNA integrity—the degradation level of a particular RNA sample has a direct impact on the percentage of mRNA target that is converted into cDNA and therefore quantified. Depending on the first-strand primer used, degradation may prevent the reverse transcriptase from creating cDNA that corresponds with full-length target amplicon. The lower the reverse transcription efficiency, the less sensitive the PCR assay will be. Efficiency variations that are not normalized can result in inaccurate conclusions.

GC content, RNA sample complexity, and reverse transcriptase employed—RNA expression level comparisons are more accurate if the RT is less sensitive to the inevitable differences between samples. For example, data have shown that sample complexity alone, meaning all the background nucleic acid not compatible with the reverse transcription primer, can result in as much as a 10-fold difference in reaction efficiency. Reverse transcriptases capable of consistent cDNA synthesis in this background are ideal.

Carryover of organic solvents and chaotropic salts—ethanol and guanidine are necessary for RNA capture but can inhibit enzymatic reactions. Variations in the levels of these contaminants between RNA samples can affect sample comparison. Therefore, it is important to use an RNA isolation method that results in consistently low levels of these byproducts. We also recommend using a validated normalizer gene for real-time PCR.

2.6 Controls

Controls in real-time PCR reactions prove that signals obtained from experimental samples represent the amplicon of interest, thereby validating specificity. All experiments should include a no-template control, and qRT-PCR reactions should also include a no-reverse transcriptase (no-RT) control.

No-template controls should contain all reaction components except the DNA or cDNA sample. Amplification detected in these wells is due to either primer-dimers or contamination with amplified PCR product. This type of contamination can make expression levels look higher than they actually are.

No-RT reactions should contain all reaction components except the reverse transcriptase. If amplification products are seen in no-RT control reactions, it indicates that DNA was amplified rather than cDNA. This can also artificially inflate apparent expression levels in experimental samples.

2.7 Normalization methods

Earlier in this guide, eliminating experimental inconsistencies was mentioned as a paramount concern for real-time PCR experimental design. Deviating from the experimental plan can limit researchers' ability to compare data and could lead to erroneous conclusions if deviations are not accounted for in the analysis. Sources of experimental variability include the nature and amount of starting sample, the RNA isolation process, reverse transcription, and, of course, real-time PCR amplification. Normalization is essentially the process of neutralizing the effects of variability from these sources. While there are individual normalization strategies at each stage of real-time PCR, some are more effective than others. These strategies include:

Normalizing to sample quantity—initiating the RNA or DNA isolation with a similar amount of sample (e.g., tissue or cells) can minimize variability, but it is only approximate and does not address biases in RNA isolation.

Normalizing to RNA or DNA quantity—precise quantification and quality assessment of the RNA or DNA samples are necessary, but fall short as the only methods for normalization, because they do not control for differences in efficiency in reverse transcription and real-time PCR reactions. For example, minute differences in the levels of contaminants can affect reverse transcription and lower amplification efficiency. Any variation in samples is then amplified during PCR, with the potential to result in drastic fold-changes unrelated to biological conditions within samples. Pipetting is also subject to operator variation, and there is no normalization to compensate for it in post-purification RNA analysis.

Normalizing to a reference gene—the use of a normalizer gene (also called a reference gene or endogenous control) is the most thorough method of addressing almost every source of variability in real-time PCR. However, for this method to work, the gene must be present at a consistent level in all samples being compared. An effective normalizer gene controls for RNA quality and quantity, and differences in both reverse transcription and real-time PCR amplification efficiencies. If the reverse transcriptase transcribes or the DNA polymerase amplifies a target at different rates in two different samples, the normalizer transcript will reflect that variability. Endogenous reference genes, such as a housekeeping gene, or exogenous nucleic acid targets can be used.

Endogenous controls

Common endogenous normalizers in real-time PCR include the genes encoding:

- β-actin: cytoskeletal component
- 18S rRNA: ribosomal subunit
- Cyclophilin A (PPIA): serine-threonine phosphatase inhibitor
- Glyceraldehyde 3-phosphate dehydrogenase (GAPDH): glycolysis pathway
- β-2-microglobulin (B2M): major histocompatibility complex
- β-glucuronidase (GUS): exoglycosidase in lysosomes
- Hypoxanthine ribosyltransferase (HPRT1): purine salvage pathway
- TATA-box binding protein (TBP): RNA transcription

Because every real-time PCR experiment is different, thought and careful planning should go into selecting a normalizer. Instead of choosing a normalizer based on what others in the lab use, choose one that best supports the quantification strategy of the specific target.

The first requirement of a quality normalizer is that it is similar in abundance to the target gene product. This is especially important when multiplexing, because if the normalizer reaction plateaus before the target C_t is reached, the normalizer itself will impede target amplification, causing a higher target C_t, thus defeating its purpose. Alternatively, the normalizer reaction can be assembled with primer-limited conditions to mimic a lower expression level. Most TaqMan endogenous control assays are available with primer-limited configuration.

Real-time PCR assays for normalizer targets should have a similar amplification efficiency to assays for experimental targets; this can be evaluated using a standard curve. Although correction factors can be applied when comparing reactions with different efficiencies, accuracy is enhanced when the reaction efficiencies are close to one another.

Last and most important, expression of the normalizer should be consistent, regardless of the treatment or disease state of the sample. This must be experimentally determined, as shown in Figure 19.

While multiple replicates should be performed to ensure accuracy, it is clear that *PPIA*, *TBP*, and *HPRT1* would not be good normalizer choices for these treatment groups

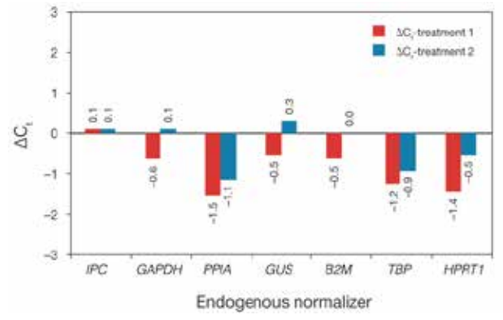


Figure 19. Gene expression levels of commonly used endogenous controls and the importance of normalization. In this example, two treatment groups and a normal group were analyzed for the expression levels of common reference genes, which were amplified in addition to an internal positive control (IPC). The IPC provides a standard for normal reaction-to-reaction variability. The bars represent up- or downregulation of the normalizer in each treatment group as compared to the normal sample, which is represented by a ΔC_t of 0. The goal is to find a normalizer that mimics the changes exhibited by the IPC.

because they appear to be downregulated with treatment. The expression levels of even the most common reference genes can be altered under certain conditions and therefore should always be validated:

- GAPDH is a common normalizer that has been shown to be consistent in many cases. However, GAPDH is upregulated in some cancerous cells, in cells treated with tumor suppressors, under hypoxic conditions, and in manganese- or insulin-treated samples.
- β-actin is another commonly employed housekeeping gene because it exhibits moderately abundant expression in most cell types. However, its consistency has been questioned in breast epithelial cells, blastomeres, porcine tissues, and canine myocardium.
- 18S rRNA constitutes 85–90% of total cellular RNA and has been shown to be quite consistent in rat liver, human skin fibroblasts, and human and mouse malignant cell lines. However, its level of abundance makes it a problematic normalizer for medium- and low-expression targets. Often it is difficult to find a concentration of RNA at which 18S rRNA provides a wide enough baseline and also at which the target of interest generates a C_t within 40 cycles. In addition, multiplexing may necessitate limiting the concentration of 18S primers so that the normalizer doesn't sequester all the reaction components and make PCR conditions unfavorable for the target of interest.

Alternative methods exist that do not rely on the accuracy of a single reference gene, but rather the geometric mean of multiple validated normalizers. This use of multiple consistent normalizers may prove to be a better buffer against the C_t fluctuations of any single gene, thereby increasing assay and sample type flexibility.

Exogenous normalizers

Exogenous normalizers are not as commonly employed, but are a viable alternative if a highly consistent endogenous normalizer cannot be found for a specific sample set. An exogenous reference gene is a synthetic or *in vitro*–transcribed RNA with a sequence that is not present in the experimental samples. Due to its exogenous origin, it does not undergo the normal biological fluctuations that can occur in a cell under different conditions or treatments. When using exogenous normalizers, the earlier they are added to the experimental workflow, the more steps they can control. For example, if an exogenous transcript is added to the cell lysis buffer, it can be used as a normalizer for cell lysis, RNA purification, and subsequent reverse transcription and PCR reactions.

An example of an exogenous normalizer is an *in vitro*–transcribed RNA specific to plant processes, such as a photosynthetic gene. This could be spiked into mammalian samples because those cells would not have this same transcript endogenously.

The drawbacks to employing an exogenous normalizer are:

- It is not endogenous; maximize the utility of exogenous normalizers by spiking them into the workflow early, for example into the cell lysis buffer
- Accuracy is subject to pipetting variability when introducing the normalizer
- Transcript stability may be affected by prolonged storage and multiple freeze-thaws; therefore, copy number should be routinely assessed to ensure it has not shifted over time

2.8 Using a standard curve to assess efficiency, sensitivity, and reproducibility

The final stage of experimental design is validating that the parameters discussed up to this point result in a highly efficient, sensitive, and reproducible experiment.

Reaction efficiency

As discussed previously, the overall efficiency of real-time PCR depends on the individual efficiencies of the reverse transcription reaction and the PCR amplification reaction. Reverse transcription efficiency is determined by the percentage of target RNA that is converted into cDNA. Low conversion rates can affect sensitivity, but variation in the conversion percentage across samples is of greater concern.

PCR amplification efficiency is the most consistent factor in a real-time PCR reaction. However, this amplification exponentially magnifies slight variations in reverse transcription efficiency, potentially resulting in misleading data. An efficiency of 100% corresponds to a perfect doubling of template at every cycle, but the acceptable range is 90–110% for assay validation. This efficiency range corresponds to standard curve slopes of –3.6 to –3.1. The graph in Figure 20 shows the measurement bias resulting solely from differences in reaction efficiency.

Validating the reaction efficiency for all targets being compared (e.g., reference genes and genes of interest), optimizing those efficiencies to be as similar as possible, and employing efficiency corrections during data analysis can reduce these effects. These strategies will be discussed further in the data analysis section.

Reaction efficiency is best assessed through the generation of a standard curve. A standard curve is generated by creating a dilution series of sample nucleic acid and performing real-time PCR. Then, results are plotted with input nucleic acid quantity on the x-axis and C_t on the y-axis. Samples used to generate the standard curve should match (as closely as possible) those that will be used for the experiment (i.e., the same total RNA or DNA sample). The dilution range, or dynamic range analyzed for the standard curve, should span the concentration range expected for the experimental samples. The slope of the curve is used to determine the reaction efficiency, which, as noted previously, most scientists agree should be between 90% and 110%.

In Figure 21, standard curves generated for three different targets are shown. The parallel nature of the red and blue curves indicates that they have similar efficiencies and therefore can be accurately compared at any dilution. An example of this type of comparison would be when a normalizer gene is compared against a target gene to adjust for non-biological variability from sample to sample. The purple curve, however, becomes less efficient at the lower concentrations and therefore cannot accurately be used for comparison purposes at these lower concentrations.

In addition to assessing the experimental conditions and providing an efficiency value for relative quantitation, standard curves can also be used to determine whether the problem with a particular reaction is due to inhibition or lack of optimization. This will be discussed in more detail in the troubleshooting section.

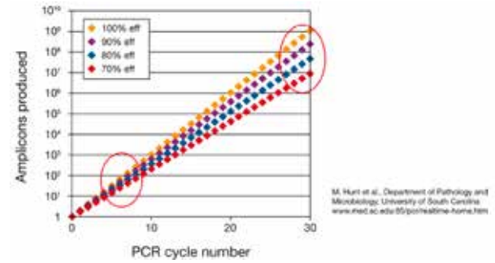


Figure 20. Bias effect caused by different amplification efficiencies. Four different real-time PCR reactions are shown that range from 70% to 100% efficiency. The divergence is not necessarily apparent in the early cycles. However, after 30 cycles, there is a 100-fold difference in reported copy number between a reaction with 70% efficiency and one with 100% efficiency. Differences in efficiency become more important in reactions with more cycles and in assays that require greater sensitivity.

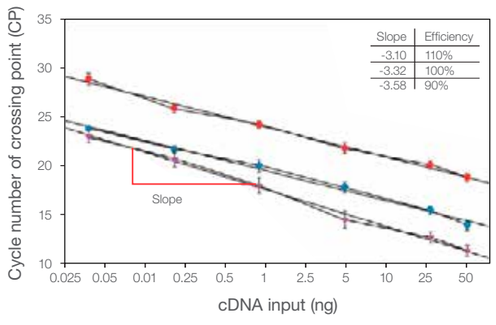


Figure 21. Example standard curve used to evaluate the efficiency of real-time PCR.

Sensitivity and reproducibility

A standard curve with an efficiency within the desirable window of 90–110% defines the range of input template quantities that may be measured in real-time PCR. To some, sensitivity is measured by how early a target C_t appears in the amplification plot. However, the true gauge of sensitivity of an assay is whether a given low amount of template fits the standard curve while maintaining a desirable amplification efficiency. The most dilute sample that fits determines reaction sensitivity.

The standard curve also includes an R^2 value, which is a measure of replicate reproducibility. Standard curves may be repeated over time to assess whether the consistency, and therefore the data accuracy for the samples, is maintained.

Plate preparation

3

Contents

3.1 Mixing

It is a good idea to briefly mix, then centrifuge all real-time PCR reaction components just before assembling reactions. Gently swirl enzyme-containing master mixes and briefly (1–2 seconds) vortex other components such as PCR primer pairs or TaqMan Assays, and thawed nucleic acid samples. Since real-time PCR master mixes are typically denser than the other real-time PCR reaction components, it is important to adequately mix reaction mixtures, because otherwise, precision could be compromised. Gently inverting tubes a few times or lightly vortexing for 1–2 seconds are highly efficient mixing methods. However, avoid overvortexing, because it can cause bubbles to form that could interfere with fluorescence detection, and can reduce enzyme activity, which could reduce amplification efficiency. Always centrifuge briefly to collect the contents at the bottom of the container and eliminate any air bubbles from the solutions.

3.2 Plate loading

Base the order of reagent addition into reaction plates or tubes on the nature of the experiment. For example, to analyze the expression of 5 genes in 20 different RNA samples, it would make more sense to dispense assay mix into the reaction plate or tube first and then add sample. On the other hand, to analyze the expression of 20 genes in only 5 different RNA samples, it would be easier to first dispense the RNA samples, then add assay master mix. Regardless of the order of reagent addition, plan your pipetting to avoid cross-contamination of samples and assays. Mixing fully assembled reactions is not necessary, because nucleic acids are hydrophilic and will quickly mix with all other reaction components.

3.3 Plate sealing

Plastic PCR plates may be sealed with optical caps or optical covers, which are thin sheets of plastic with adhesive on one side. The adhesive is protected by a white backing. On the ends of the cover are rectangular tabs delineated with perforations. Use these tabs to handle the cover without touching the cover itself. Remove the white backing and place the cover on top of the plate, inside the raised edges. A square plastic installing tool is provided with the covers. Use it to smooth down the cover, especially along the four edges at the top of the plate. The edge of the installing tool may also be used to hold down each end of the cover as the white strips are removed by the perforation, so as not to pull up the cover. Inspect the top of the plate to verify that the cover is in good contact with the plate, especially along the four upper edges.

Once the plate has been sealed, hold it up and inspect the bottom, looking for any anomalies, such as unintended empty wells, drops of liquid adhering to the walls of the well, and air bubbles at the bottom of the well. Note any empty wells or wells with abnormal volumes. Centrifuge the plate briefly to correct any adherent drop and bottom-bubble problems.

Plates may be safely labeled along the skirt. Alternatively, plates are available with bar codes. Do not write on the surface of the plate over well positions, as this will interefere with fluorescence excitation and reading. Also, do not write on the bottom of plates, as the ink could be transferred to the block in the real-time PCR instrument. Significant block contamination by colored or fluorescent substances can adversely affect real-time PCR data.

3.4 Plate insertion

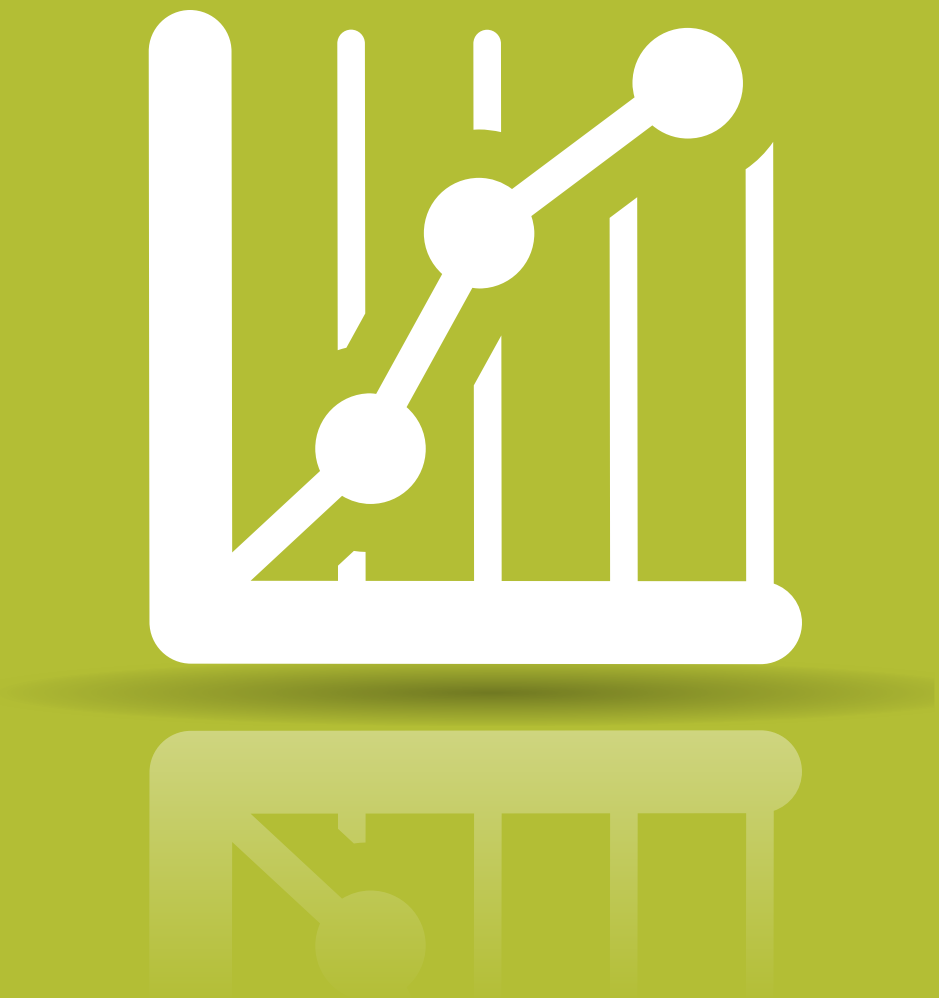
Once the plate is ready, it may be loaded into the real-time PCR instrument, following the manufacturer's instructions. Plates containing DNA or cDNA template and Applied Biosystems™ AmpliTaq Gold™ reagents are very stable at ambient temperature. They can be routinely stored at room temperature under normal laboratory lighting for days without ill effect. Avoid direct exposure of loaded plates to sunlight; this can compromise the fluorescent dyes in the mixture. If a plate needs to be transported outside, wrap it in aluminum foil to protect it from sunlight.

For instruments with interchangeable thermal cycling blocks, be sure that the block matching the plate type is installed. Note that a 96-well standard block and 96-well Fast block are not the same. Standard 96-well plates and standard tubes have a 0.2 mL capacity, whereas Fast 96-well plates and Fast tubes can accommodate 0.1 mL. Fast plastics must be used with a Fast block, even if Fast mode is not being used.

The loading process varies with the model of instrument. Some instruments have a drawer system: the drawer is pulled out, the plate is placed in the block or holder, and the drawer is pushed back in. For some instruments, the plate is placed on an arm, which is computer controlled.

Data analysis

4



Contents

4.1 Introduction

As mentioned in the beginning of the experimental design section, selecting the right quantification method depends on the goals of the experiment.

- Absolute quantification determines actual copy numbers of target, but is also the most labor-intensive and difficult form of quantitation. This method requires thoughtful planning and a highly accurate standard curve. Absolute quantitation is often used for determining viral titer.
- Comparative quantification still requires careful planning, but the data generated are for relative abundance rather than exact copy number. This is the method of choice for gene expression studies, and it offers two main options for quantification: $\Delta\Delta C_t$ and standard curve quantification.

4.2 Absolute quantification

Absolute quantification is the real-time PCR analysis of choice for researchers who need to determine the actual copy number of the target under investigation. To perform absolute quantification, a target template solution of known concentration is diluted over several orders of magnitude, amplified by real-time PCR, and the data are used to generate a standard curve in which each target concentration is plotted against the resulting C_t value. The unknown sample C_t values are then compared to this standard curve to determine their copy number.

Standard curve generation—overview

With an absolute standard curve, the copy number of the target of interest must be known. This dictates that, prior to curve generation, the template must be accurately quantified. Figure 22 highlights a standard curve setup. A sample of the target of interest is accurately determined to contain 2 x 10^{11} copies. The sample is diluted 10-fold eight times, down to 2 x 10^3 copies, and real-time PCR is performed on each dilution, using at least three replicates. The resulting standard curve correlates each copy number with a particular C_t. The copy number values for the unknown samples are then derived by comparison to this standard curve. The accuracy of the quantification is directly related to the quality of the standard curve.

In absolute quantification, consider the following:

1. The template for standard curve generation and the method used to quantify that template are the foundations for the experiment. Pipetting accuracy for the dilution series is essential. Also remember that real-time PCR sensitivity amplifies minute human error.
2. Similar reverse transcription and PCR efficiencies for the target template and dilution series of the actual samples are critical.

Template choice for standard curve generation

As mentioned, the template used for absolute standard curve generation will determine the accuracy of the data. Although a homogeneous, pure template may be needed for initial copy number determination, for generation of the standard curve, it is best to use target template that is as similar to the experimental samples as possible. Because steps such as nucleic acid isolation and reverse transcription play a role in reaction dynamics, this includes subjecting it to most of the same processing steps as the experimental samples. The following types of template have been used as absolute quantification standards:

1. **DNA standards**—PCR amplicon of the target of interest, or plasmid clone containing the target of interest.

Pros: Easy to generate, quantify, and maintain stability with proper storage

Cons: Cannot undergo the reverse transcription step of qRT-PCR, which can impact reaction efficiency significantly

2. **RNA standards**—*in vitro*–transcribed RNA of the target of interest (Figure 23).

Pros: Incorporates reverse transcriptase efficiency and mimics the target of interest most similarly

Cons: Time-consuming to generate and difficult to maintain accuracy over time due to instability

The homogeneous nature of each of the RNA and DNA standards means that they will often exhibit higher efficiencies than experimental samples. Therefore, background RNA, such as yeast tRNA, can be spiked into the standard template to create a more realistic heterogeneous environment and help to balance reaction efficiency. It has been shown that background RNA can suppress the cDNA synthesis rate as much as 10-fold.

Standard curve application—complementary RNA (cRNA)

To demonstrate how the recommendations for an absolute quantification standard curve should be applied, this section will walk step by step through the creation of a cRNA standard curve for this method of quantification.

T7 RNA polymerase can be used to generate a homogeneous pool of the transcript of interest from a plasmid or a PCR product.

Because it has an extended limit of detection and better accuracy, fluorometric measurement of cRNA is recommended over UV absorbance measurement. Another precise method of quantification is digital PCR.

With the copy number determined, unrelated yeast tRNA can be added at a 1:100 cRNA to tRNA ratio to mimic the normal background of biological samples. This standard is then diluted over at least 5 to 6 orders of magnitude for use in C_t determination by real-time PCR.

Pipetting inaccuracies can have a significant effect on absolute quantification data. Appropriate precautions can minimize this effect. As can be seen in the plate setup in Figure 24, three separate cRNA dilution series are prepared, and each dilution within each series is amplified in duplicate. It is important that the dilution series encompass all possible template quantities that may be encountered in the experimental samples. For example, the lowest point in the standard curve should not contain 100 template copies if it is possible that an unknown test sample may contain only 10 copies.

For each dilution, six C_t values will be obtained. The high and low C_t values are discarded, and the remaining four C_t values are averaged. Focusing on the 10^{-4} dilution in this example (Figure 24), it can be seen how the C_t values for a given sample vary by as many as 2 cycles. This is minimized by assigning this dilution, which corresponds to a particular copy number, the average C_t value of 21.4.

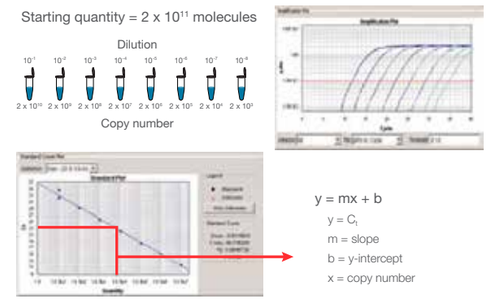


Figure 22. Workflow for standard curve setup for absolute quantification.

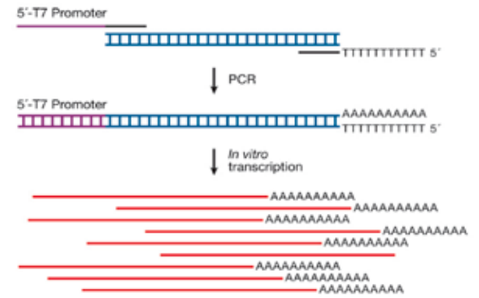


Figure 23. Schematic diagram of the *in vitro* transcription protocol. The PCR product generated from the real-time PCR itself can be reamplified with a 5′ T7 promoter-containing sequence and a 3′ poly(T)-containing reverse primer. The *in vitro* transcription reaction produces polyadenylated sense mRNA. After purification, it will be accurately quantified and diluted for the standard curve.

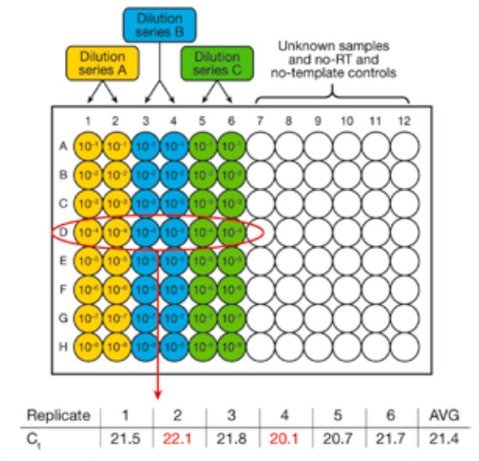


Figure 24. Plate setup for standard curve generation. When calculating the average C_t value, the highest and lowest values obtained from replicate reactions can be discarded.

4.3 Comparative quantification

Comparative quantification, while still technically challenging, is not quite as rigorous as absolute quantification. In this technique, which applies to most gene expression studies, the expression level of a gene of interest is assayed for up- or downregulation in a calibrator (normal) sample and one or more experimental samples. Precise copy number determination is not necessary with this technique, which instead focuses on fold change compared to the calibrator sample.

Here the common methods of comparative quantification and how variability is controlled in each are outlined.

Comparative quantification algorithms—ΔC_t

This is comparative quantification in its most basic form. C_t values are obtained for expression of the gene of interest from both a test and calibrator sample, and the difference between them is the ΔC_t. The fold difference is then simply 2 to the power of ΔC_t.

Fold difference = $2^{\Delta Ct}$

This basic method is inadequate because it does not control for differences in sample quantity, sample quality, or reaction efficiency (Figure 25).

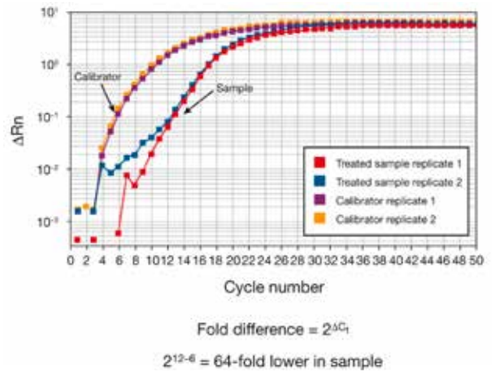


Figure 25. Comparative quantification of expression in a treated sample and calibrator. The treated sample and calibrator were run in duplicate; the calibrator C_t was 6 and it was 12 for the treated sample. According to the ΔC_t method, the calculated relative expression level of the target gene in the treated sample is 64-fold lower than that of the calibrator. However, because a normalizer was not employed, the effects of experimental variability on this result are unknown, and thus the conclusion cannot be considered trustworthy.

Comparative quantification algorithms—$\Delta\Delta C_t$

The $\Delta\Delta C_t$ method is a very popular technique that compares results from experimental samples with both a calibrator (e.g., untreated or wild type sample) and a normalizer (e.g., housekeeping gene). With this method, C_t values for the gene of interest (GOI) in both the test sample(s) and calibrator sample are now adjusted in relation to a normalizer (norm) gene C_t from the same two samples. The resulting $\Delta\Delta C_t$ value is incorporated to determine the fold difference in expression.

Fold difference = $2^{-\Delta\Delta Ct}$

$\Delta C_{t\ sample} - \Delta C_{t\ calibrator} = \Delta\Delta C_t$

$C_{t\ GOI}{}^{s} - C_{t\ norm}{}^{s} = \Delta C_{t\ sample}$

$C_{t\ GOI}{}^{c} - C_{t\ norm}{}^{c} = \Delta C_{t\ calibrator}$

The requirement for the $\Delta\Delta C_t$ method is that the efficiencies for both the normalizer and target gene are identical. Of course, the obvious question is: what range of deviation is acceptable? The way to determine this is to generate a standard curve for both the normalizer gene and target gene of interest using the same samples (Figure 26). The average ΔC_t between the normalizer and target gene can be obtained for each dilution. The value itself is not important; it is the consistency of that value across each dilution that matters.

To some researchers, this small deviation in efficiency still opens the door to inaccuracies. Employing a correction for both the gene of interest and the normalizer minimizes the effects of variation in amplification efficiency.

Comparative quantification algorithms—standard curve method

The standard curve method of comparative quantification employs the C_t difference between the target gene in the test and calibrator samples, normalized to the reference gene C_t values and adjusted for minute variations in amplification efficiency. A standard curve to determine amplification efficiency for both the normalizer and the gene of interest is necessary with this method, but it does avoid the assumptions made with previous techniques. A requirement of this technique is that the normalizer gene be the same across all samples in the analysis.

In Figure 27,

Fold difference = $(E_{target})^{\Delta Ct\ target} / (E_{normalizer})^{\Delta Ct\ normalizer}$

E = efficiency from standard curve $E = 10^{[-1/slope]}$

$\Delta C_{t\ target} = C_{t\ GOI}{}^{c} - C_{t\ GOI}{}^{s}$

$\Delta C_{t\ normalizer} = C_{t\ norm}{}^{c} - C_{t\ norm}{}^{s}$

Fold difference equation derived from M.W. Pfaffl in A-Z of Quantitative PCR.

In order to capture the most accurate efficiency value for the calculation, choose the calibrator sample for the standard curve carefully. This calibrator sample should undergo the same purification procedure, be involved in the same reactions, and have similar complexity to the experimental samples. Therefore, the perfect calibrator sample is one of the heterogeneous samples containing the target of interest—for example, total RNA from the cell line or tissue being studied. Keep in mind that any differences between the calibrator sample and the experimental samples may result in an inaccurate efficiency correction and therefore inaccurate calculations for fold changes in gene expression. Because copy number is irrelevant, the dilutions or values given to those dilutions can be arbitrary.

While the same setup for the amplification curves from the $\Delta\Delta C_t$ method are used in this technique, efficiency values from standard curves (ideally run on the same plate) are now incorporated to adjust the normalizer and gene-of-interest C_t values. Efficiencies are derived from both slopes (Figure 27).

In summary, the first step in choosing a quantification strategy is to determine whether absolute or relative quantification will best address the questions to be answered. To find out how many target molecules are in the sample, a precise standard curve using known quantities of target template must be generated. In most cases, relative quantification will be the method of choice. The ΔC_t method does not employ a normalizer, while the $\Delta\Delta C_t$ method involves one or more reference genes to normalize for real-time PCR processing variability. Employing a normalizer gives the option of fold-change calculations, with or without an adjustment in reaction efficiency.

Input amount RNA (ng)	Normalizer gene Avg C_t	Target gene Avg C_t	ΔC_t Normalizer - target
10	20.3	18.2	2.1
5	21.4	19.4	2
1	22.8	20.6	2.2
0.5	23.8	21.3	2.5
0.1	24.8	22.5	2.3
0.01	25.9	23.7	2.2
0.05	26.8	24.5	2.3

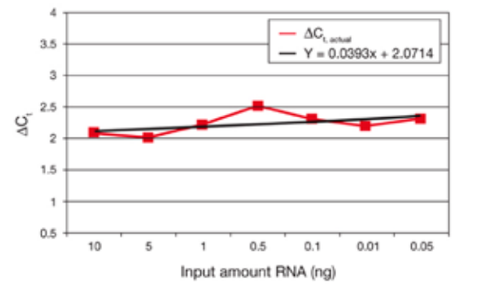


Figure 26. Relative efficiency plot. The ΔC_t ranges from 2.0 to 2.5. When plotted against dilution or input amount of RNA, a slope is obtained. While a perfectly flat line (slope = 0) indicates identical efficiency across all input concentrations, a slope of <0.1 is generally considered acceptable when employing the $\Delta\Delta C_t$ method.

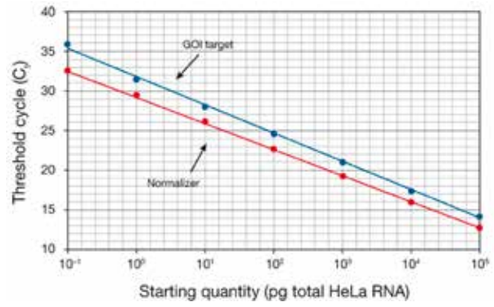


Figure 27. Efficiency values are obtained from standard curves to adjust the normalizer and gene-of-interest C_t values. HeLa RNA was diluted over 6 orders of magnitude, and real-time PCR was performed to generate standard curves for both the normalizer and gene of interest (GOI).

4.4 High resolution melting curve analysis

High resolution melting curve analysis is a homogeneous, closed-tube post-PCR method for identifying SNPs, novel mutations, and methylation patterns. High resolution melting curve analysis is a more sensitive approach than traditional melting curve profiling, in which dsDNA is monitored for the temperature at which it dissociates into ssDNA, its T_m. High resolution melting curve profiling requires a real-time PCR instrument with upgraded optical and thermal capabilities as well as analysis software for extremely fast data acquisition and highly accurate thermal control and consistency. The Applied Biosystems™ QuantStudio™ 3, 5, 6 Flex, 7 Flex, 12K Flex, ViiA™ 7, 7900HT Fast, 7500 Fast, StepOnePlus™, and StepOne™ real-time PCR systems; the Qiagen™ Rotor-Gene™ 6000 System; and the Roche Diagnostics™ LightCycler™ 480 System are commercially available instruments currently set up to perform high resolution melting curve analysis.

In this type of analysis, approximately 80–250 bp fragments of genes are amplified using PCR in mixtures that contain a high-performance dsDNA-binding dye. The Applied Biosystems™ MeltDoctor™ HRM Master Mix employs MeltDoctor HRM Dye, a thermostabilized form of the fluorescent SYTO™ 9 dye with low background fluorescence and high brightness in the presence of dsDNA. After 40 cycles, the amplification products are annealed and highly fluorescent. When the high resolution melting curve analysis begins, the real-time PCR instrument slowly ramps the temperature higher while simultaneously recording fluorescence data from the amplicons. As the PCR products begin to denature (or melt), their fluorescence will slowly decrease due to the fluorescent dye being released, until the temperature approaches the PCR product's T_m. Very close to the T_m, a dramatic decrease in fluorescence is observed as the sample transitions from dsDNA to ssDNA (Figure 28).

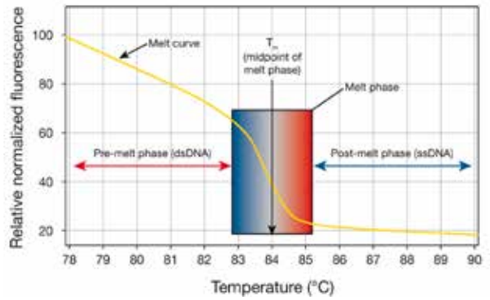


Figure 28. Characteristics of the melt-curve profile for a PCR amplicon. Specific DNA sequences have characteristic melt-curve profiles. Mutations can be detected as either a shift in T_m or a change in shape of the melting curve. In contrast to traditional melting curve analysis, high resolution melting curve analysis can distinguish between amplicons with just a single-nucleotide difference. This technique has opened the door to many new applications for dsDNA-binding dyes.

High resolution melting curve applications

The most widely used high resolution melting curve application is gene scanning. Gene scanning is the search for the presence of unknown variations in PCR amplicons prior to, or as an alternative to, sequencing. Mutations in PCR products are detectable by high resolution melting curve analysis because they lead to changes in the shape of DNA melting curves. When amplified and melted, heteroduplex DNA samples show melting curves with different profiles than those derived from homozygous wild type or mutant samples.

Some common applications of high resolution melting curve analysis include:

- Gene scanning (mutation discovery)
- Mutation analysis
- SNP detection
- Heterozygosity studies
- Species identification
- Methylation analysis

These applications traditionally required a unique fluorogenic probe for each target, which was expensive, time-consuming to design, and inflexible. High resolution melting curve analysis using dsDNA-binding dyes offers the same capabilities as probe-based analysis in a more inexpensive and flexible format.

While many applications exist for high resolution melting curve analysis, discrimination at the single-nucleotide level is one of the more challenging, due to the minute T_m shifts that must be detected.

High resolution melting curve analysis chemistry

In addition to a specialized instrument and software, high resolution melting curve analysis requires dsDNA-binding dyes that are capable of distinguishing the melting points of amplicons that differ by a single nucleotide.

Dyes that have been successfully used for high resolution melting curve analysis include:

- MeltDoctor HRM Dye
- SYTO 9 dye
- BioFire™ Defense LCGreen™ and LCGreen™ Plus+ reagents
- Biotium™ EvaGreen™ dye

Keys to successful high resolution melting curve assays

- Set an appropriate ramp rate for the instrument. This will typically result in 10–20 data collection points per °C, which is needed for high resolution melting curve analysis.
- Keep amplicons short for highest sensitivity. Compared to larger amplicons, those around 100 bp will allow easier detection of single-nucleotide melt events.
- Ensure specificity of the PCR amplicon. Mispriming products and primer-dimers can complicate data interpretation. Primer concentrations lower than 200 nM, $MgCl_2$ in the 1.5 mM to 3 mM range, and use of a hot-start DNA polymerase will help to obtain high specificity. Assess mispriming using a standard low resolution melting curve. No-template control melting curves are important for evaluating specificity.
- Avoid amplifying across regions that could contain variations other than those of interest. Check for species homology, exon/intron boundaries, splice sites, and secondary structure and folding of the PCR product.
- Maintain similar fluorescent plateaus, and therefore similar PCR product quantities, for all targets being analyzed. Differences in quantity between the samples being compared can affect melting temperatures and confound analysis. Starting with similar amounts of template can be helpful.
- Ensure that enough template is used in the reaction. In general, C_t values should be below 30 to generate sufficient material for accurate melting curve analysis.
- Provide a sufficient melt data collection window. For example, the window should have a range of 10°C on either side of the melting temperature of the amplicon (T_m ± 10°C). Sufficient pre- and post-melting temperature data are required for accurate curve normalization and high replicate correlation.
- It is also recommended, for some instruments, to insert a pre-hold step at 50°C following amplification (but prior to melting) to ensure that all products have reassociated and to encourage heteroduplex formation. Use a mix, such as MeltDoctor HRM Master Mix, that offers sensitive and unbiased amplification of the gene fragments of interest.

4.5 Multiplex real-time PCR analysis

Multiplexing is a technique in which more than one target is analyzed in the same real-time PCR reaction. Each target is distinguished by a particular dye, conjugated to the fluorogenic probe or primer pair specific for that target. Typically, multiplex reactions are used to amplify a normalizer gene and a gene of interest in the same reaction.

Theoretically, the number of targets that can be amplified in a given reaction is limited only by the number of available spectrally distinct dyes and the number of dyes that can be excited and detected by the real-time PCR instrument. However, other experimental hurdles exist: the different primer pairs and/or probes in a reaction must not interact with each other, and these primers and probes have to share available PCR components, such as dNTPs and thermostable DNA polymerase. While more time-consuming to optimize, multiplexing offers several distinct advantages:

- **Less variability, more consistency**—multiplexing a normalizer and gene of interest in the same tube eliminates well-to-well variability that could arise if those same two targets were amplified in different (even adjacent) wells
- **Less reagent usage, less cost**—multiplexing requires fewer reactions to analyze the same number of targets
- **Higher throughput**—more targets can be analyzed per real-time PCR run and per sample

Keys to a successful multiplex assay

A successful multiplex reaction must take many factors into consideration, including:

- Primer and probe design
- Reagent optimization (including primer concentration, target abundance, reaction components, and fluorophore/quencher combinations)
- Validation of the multiplex assay

Primer and probe design

Primer and probe design is arguably the most critical factor in a multiplex assay. As reaction complexity increases, so does the probability that primers and probes will dimerize or that competition for reaction components will limit the amplification of one or more targets. The following are particularly important factors that will maximize performance while minimizing competitive effects:

- Keep amplicons short. Designing primers to amplify a segment ranging from 60 bp to 150 bp will enhance reaction efficiency.
- Design primers with T_m values within 1°C of each another. Remember, all primers and probes will be annealing at one temperature. Mismatched T_m values will result in an efficiency bias.
- Perform BLAST searches with primer and probe designs to ensure their specificity for the targets of interest.
- Use primer design software to determine whether any of the primer or probe sequences are prone to dimerization. One such free program, AutoDimer (authored by P.M. Vallone, National Institute of Standards and Technology, USA), can analyze all the primer designs in a given multiplex reaction for their propensity to dimerize in any combination. Another program, Multiple Primer Analyzer, is available on the Thermo Fisher Scientific website.

Reagent optimization

A critical concern in multiplex reactions is the competition for reagents among the different targets being amplified. Ensuring a high efficiency of amplification requires reducing the amount of primers, increasing the concentrations of all components, or both. Most optimization experiments begin by testing different primer concentrations.

Primer concentration and target abundance

Not all amplicons in a multiplex reaction are present in the same numbers nor amplify at the same efficiency. Limiting the primer concentration for those targets that are easier to amplify or more abundant can allow more consistent amplification across targets. For example, if β-actin (a high-copy normalizer) is multiplexed with a low-abundance target, it may exhaust the shared reaction components in the early cycles. Reducing the amount of primers for β-actin limits its rate of amplification, allowing the less-abundant target to be amplified in an unhindered manner. In general, for higher-abundance targets, use the lowest primer concentration that does not increase C_t values.

Reaction component concentrations

Every multiplexing reaction is different, but adjusting primer concentrations is most likely to result in C_t values that are consistent between singleplex and multiplex reactions. Other alternatives include increasing the amounts of *Taq* DNA polymerase, magnesium, and dNTPs and increasing the buffer strength to try to boost the sensitivity and amplification efficiency of all targets involved.

Fluorophore/quencher combinations

The reporter fluorophores in a multiplex reaction must be spectrally distinct so that the fluorescence signal arising from each is detected in a single channel. With compatible dyes, the real-time PCR instrument excitation and emission optics are able to filter the wavelengths so that little, if any, fluorescence is attributed to the wrong fluorophore (such interference is also called cross-talk).

Similarly, the choice of quencher for each dual-labeled fluorescent probe becomes more important as the number of probes being multiplexed increases. Fluorescent quenchers such as TAMRA™ dye, a common quencher for FAM™ dye, work by releasing the energy of the fluorophore at a different wavelength. In a multiplex reaction, quenchers of this type result in multiple signals at different wavelengths, which can complicate filtering and possibly data integrity. Dark quenchers such as nonfluorescent quenchers or QSY™ quenchers, on the other hand, release energy in the form of heat rather than fluorescence, and therefore keep the overall fluorescent background lower. Choose appropriate multiplex dye combinations based on the detection capabilities of the instrument being used.

Validating the multiplex assay

As with singleplex real-time PCR assays, a standard curve should be completed to assess the reaction efficiencies of all the targets in a multiplex reaction prior to running the assay. There are two main stages in this validation process:

1. Validating the primers and/or probes for each target and determining their individual efficiencies. This is the initial assessment of primer and probe design.
2. Optimizing the efficiency of the overall multiplex assay.

Evaluate each primer and probe set on an individual basis to determine the designs and conditions that are ideal for the target. This is accomplished through a dilution series standard curve, again trying to achieve close to 100% efficiency. After each primer/probe combination has been functionally validated, move on to multiplex optimization. It is important to keep in mind that the singleplex conditions for primer/probe concentrations may not be optimal when the targets are multiplexed. When building a multiplex panel, it is also advisable to add targets one at a time rather than combining all targets in the first experiment, if possible.

Using the same standard curve methodology, combine all primers and probes and perform the multiplex reaction for each dilution. Compare the resulting efficiency for each target with its corresponding singleplex real-time PCR reaction efficiency. Ideally, there will be little change between single and multiplex reaction standard curves. If the reaction efficiencies for the multiplexed targets vary by greater than 5% or fall outside the desirable range of 90–110%, optimization of the primer/probe concentrations or other components in the multiplex assay will be required.

The C_t values in the multiplexed reaction should be comparable to those obtained in the singleplex reactions so that the sensitivity is not compromised.

Troubleshooting

5

Contents

5.1 Introduction

Part of implementing an ideal real-time PCR assay involves optimization to ensure that all parameters of the reaction are finely tuned for accurate results. Assay validation accompanied by any required reaction adjustments should be performed any time a new gene is being studied or assay parameters are altered. This could involve adjusting primer concentrations and thermocycling temperatures and times, and then confirming the parameters through a standard curve assessment.

While optimization does take time, it is time well spent. The resulting assay will have the highest sensitivity and dynamic range, a high efficiency (which correlates with high accuracy), and excellent reproducibility. These factors all lead to confidence in the data and ultimately to results accepted by the research community.

Troubleshooting a real-time PCR reaction can seem daunting. However, assuming proper assay design was taken into consideration, common real-time PCR difficulties can be grouped into four main areas:

- Formation of primer-dimers
- Storage of primers and probes
- Real-time PCR inhibition and poor reaction efficiency
- Software analysis settings

Formation of primer-dimers

Primer-dimers form when partial sequence homology exists between the members of the primer pair. If the primers anneal to each other during PCR, the *Taq* DNA polymerase may be able to extend them and create a product larger than the original primers and one that is more prone to anneal erroneously as cycling progresses. Depending on its length, it is also possible for a primer to fold upon itself and therefore set up a competitive environment with the template. Reaction complexity, especially that present in multiplex reactions, increases the opportunity for these unwanted interactions to occur.

Primer-dimer formation is one of the most common problems that requires troubleshooting during real-time PCR design and validation, but many opportunities exist to eliminate them from real-time PCR reactions. Before discussing how this is done, it is first important to understand why dimer formation should be minimized or reduced.

Problems caused by primer-dimers

The effect that primer-dimers can have on a reaction depends largely on the chemistry being employed. Fluorogenic probe–based reactions tend not to be influenced as much by primer-dimers because a probe annealing and being cleaved in a primer-dimer region is an extremely rare event. In this case, competition for primers is the main factor for consideration. Reactions that rely on dsDNA-binding dyes, on the other hand, are highly dependent upon primer-dimers being absent, because the dye would bind to them nonspecifically and therefore contribute to fluorescence signal being monitored during the reaction. This in turn shifts the C_t and skews results.

While the extraneous signal is the most important factor, competition within a reaction well also has a direct impact on reaction efficiency. As mentioned earlier, poor reaction efficiency shrinks the dynamic range, which in turn decreases sensitivity.

It is best to take simple precautions during primer design to avoid dimerization in the first place. There are several free tools available to assist with this. One such tool is AutoDimer (authored by P.M. Vallone, National Institute of Standards and Technology, USA), which analyzes the sequences of primer pairs and flags those that theoretically have a tendency to dimerize.

Although bioinformatics-based primer design can reduce the likelihood of dimer formation, it is still necessary to monitor dimerization experimentally.

Determining if primer-dimers are present

Gel electrophoresis is a great way to visualize primer-dimers. Primer-dimers appear as diffuse bands near the bottom of the gel, usually below 100 bp (Figure 29). During PCR, there is competition between dimer formation and template annealing and extension. As template decreases, primer-dimers often increase.

The downside to gel analysis as the sole method of validation is that its sensitivity is in the low nanogram range and therefore may be inconclusive. The advantage of gel analysis is that the size of the product can help in the overall interpretation when melting curve data are also available.

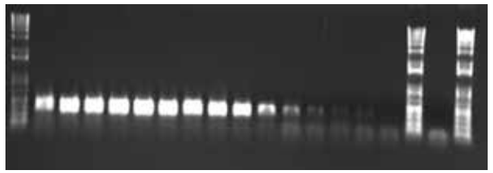


Figure 29. Agarose gel analysis to investigate primer-dimer formation. Prior to thermal cycling, the nucleic acid sample was serially diluted and added to the components of a PCR mix, and the same volume from each mixture was loaded on an agarose gel. Dimers appear as diffuse bands at the bottom of the gel.

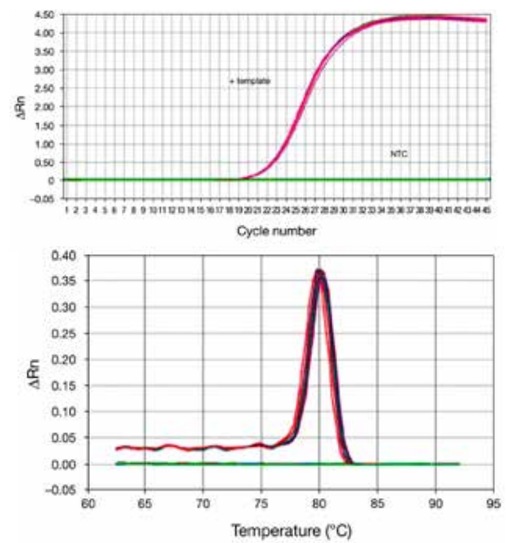
Melting curves, also referred to as dissociation curves, are a standard part of reaction thermal profiles that employ dsDNA-binding dyes. A very specific amplification will result in a single, tight peak on the melting curve for each well on the real-time PCR plate. Primer-dimers manifest themselves as lower-fluorescence, broader "waves" that indicate melting in the 70°C range. Peak shape and melting temperature are attributable to the small and variable sizes of primer-dimers. As mentioned in the data analysis section, if there is any doubt whether primer-dimers are present in a melting curve, compare the observation with the no-template control well. A primer-dimer peak is much more common when template is absent (Figure 30).

Reducing or removing primer-dimers

If primer-dimers are a concern, there are many options for reducing or eliminating their occurrence in a reaction.

1. The first is optimization of the thermal cycling conditions, which mainly involves raising the annealing temperature. In most cases, primers are designed so that they anneal successfully at 60°C, because two-step cycling (a 95°C denaturation step that goes directly to a 60°C annealing and extension step) facilitates robust amplification.
2. Primer concentration can always be lowered, and even employing different ratios of forward primer to reverse primer may help. In most cases, a final concentration of 200 nM per primer is ideal, but this can be reduced to 60 nM if necessary.
3. Magnesium is usually best at a concentration of about 3 mM. Primer-dimers are favored at concentrations above this.

A



B

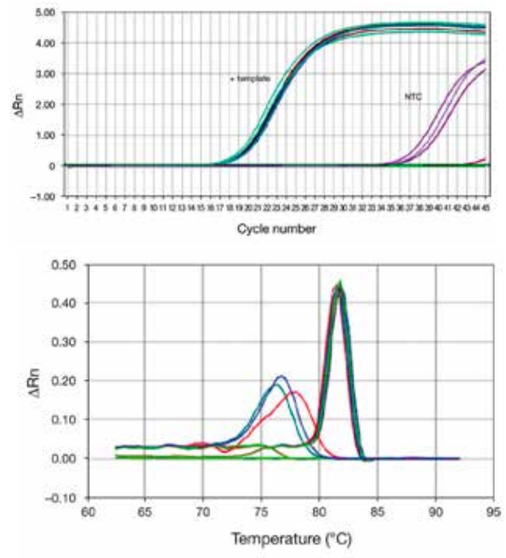


Figure 30. Amplification plots and melting profiles highlighting (A) specific amplification and (B) primer-dimer effects. The primer-dimer can be identified in (B) by the signal produced by the no-template control sample in the amplification plot, and by additional peaks in the melting profile.

4. If primers were not evaluated for their propensity toward dimerization, evaluate them and consider redesigning if necessary. And, as usual, hot-start DNA polymerases and reaction setup on ice are also preferable.
5. Ideally, more than one primer set for the same target should be tested concurrently. This can actually save much time that would otherwise be spent on optimization if one of the pairs works immediately.

Keep in mind that dimers may be more of a concern in one-step qRT-PCR due to the lower temperature of the reverse transcriptase reaction in the presence of the primer pair. Also note that there are times when primer-dimers are not a concern; for example, if they only appear in the no-template control wells, or if the ΔC_t between the samples and the no-template control well is >10 (since this means that the contribution of any fluorescence from the primer-dimers to the overall signal is negligible).

Storage of primers and probes

Although often overlooked, primer and probe storage can have a major effect on the long-term success and consistency of a real-time PCR assay. The main factors that affect primer and probe stability are the storage temperature, length of time in storage, whether they have undergone prolonged exposure to light, the concentration of the stored primer or probe, and the composition of the storage solution.

Problems caused by poor storage of primers and probes

Improper storage of primers and probes can cause them to degrade and lose specificity, which in turn affects the reaction efficiency. In assays that rely on fluorescently labeled primers and probes, a degraded probe releases free dye, which increases background and decreases the signal-to-noise ratio. This can manifest as very rough amplification curves due to the low fluorescence. Fluorescent dyes attached to primers and probes can also undergo photobleaching over time, making them less detectable in the real-time PCR instrument.

Determining if primer or probe integrity is compromised

The first preventive measure to ensure primer and probe stability is simple monitoring of the storage time. In many cases, primers and probes are stable for up to a year (or more). However, under suboptimal conditions, storage-related effects may be observed within 6 months.

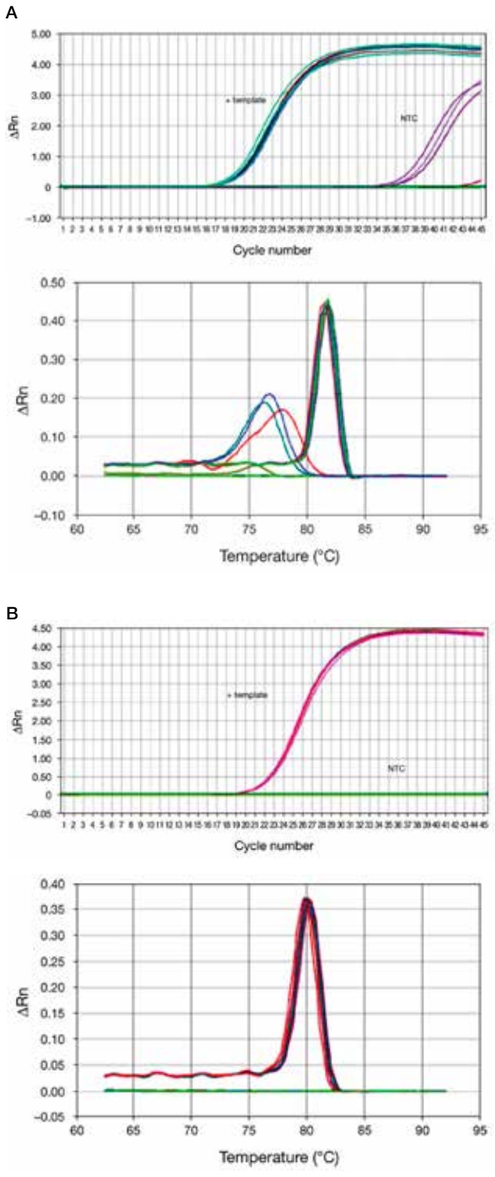


Figure 31. Amplification plots showing effects of (A) poorly stored primers (shown by the melting curve and background fluorescence) vs. (B) properly stored primers.

The best method to evaluate primer integrity is consistent employment of standard curves. Replicate inaccuracy and multiple peaks in the dissociation curve, especially if not seen previously, are common signs that stability is low.

In the case of fluorescently labeled probes and primers, observing a higher-than-normal level of background fluorescence on the instrument's multicomponent view is indicative of probe degradation.

If the fluorescent probe or primer is not degraded but the dye itself is, an ethidium bromide–stained gel can show when product is made, but not detected by the real-time PCR instrument.

Figure 31 shows standard curves highlighting the effects of poorly stored primers. The amplification plot in Figure 31A is negatively affected by degraded primers, while the curves in Figure 31B are what would be expected from properly stored primers. The melting curve provides additional detail, showing that multiple nonspecific products are present.

Maintaining primer and probe stability over time

There are four keys to maintaining primer and probe stability. Lyophilized primers have more flexibility with respect to storage time and temperature. Once reconstituted, primers should be kept at –20°C and should be monitored for signs of decreased functionality beyond a year or so. For labeled primers and probes, measures that protect the labels from light (such as the use of opaque tubes and dark storage) extend their life.

Lastly, primer concentration can have an effect on stability. Storing primers at a concentration below 10 µM is not recommended; in fact, primer concentrations of 100 µM are easier to work with in most cases. Primers and probes should also be stored in aliquots to minimize freeze-thaw cycles, especially when labeled. Tris-EDTA (TE) buffer creates a more stable environment than water. Moreover, TE that contains 0.1 mM EDTA (compared to 1 mM EDTA in standard TE) is a good choice because of the sensitivity of some PCR reactions to EDTA that may be carried over.

No-template control amplification

As discussed in the previous section, amplification signal can be observed in no-template control reactions that use dsDNA-binding dyes when primer-dimers form. But there is another case in which amplification may be seen in no-template control wells. For both probe-based and dsDNA-binding dye–based reactions, late amplification can be due to contamination. This can be a random event in which not every no-template control will show amplification, and thus this is often due to pipetting errors. If amplification products are seen in every no-template control well, then it is likely that one or more of the reagents has become contaminated. Here are some steps to take to prevent or remove contamination:

1. Use clean work spaces, including wiping down surfaces and reagents with nucleic acid–degrading solutions as needed.
2. To avoid contamination from previous PCR reactions, use master mixes containing dUTP and UDG, so that PCR products from previous reactions are degraded.
3. To determine which reagent is problematic, swap out the reagent with a new tube or different source when possible.
4. When possible, set up reactions in a different location, especially when plasmid controls are in use (which can be very easily spread but hard to remove).

Real-time PCR inhibition and poor reaction efficiency

At this point, the importance of reaction efficiency should be well understood among the critical factors in real-time PCR assay design and optimization. To review, a standard curve (generated from a dilution series of the target template) is used to obtain an efficiency value. This efficiency value acts as a marker of overall reaction "health." Low efficiency leads to one diagnosis, and high efficiency, to a different diagnosis. Steps to improve these scenarios will be quite different. While the ideal reaction efficiency is 100%, the widely accepted range is 90–110%.

Causes of high or low efficiency

Inhibition occurring in the reaction is indicated by an efficiency above 110%. Causes of inhibition include poor RNA or DNA quality, high template concentration, and carryover from nucleic acid purification. For example, if silica columns are employed, chaotropic salts used to bind the DNA or RNA might inhibit the *Taq* DNA polymerase. If organic extraction is used, phenol and ethanol carryover would have the same effect.

Poor reaction efficiency, which is an efficiency below 90%, is normally more common than inhibition. Causes include suboptimal reagent concentrations (mainly primers, magnesium, and *Taq* DNA polymerase, especially for multiplex experiments). Other factors contributing to poor reaction efficiency include primer T_m values being

more than 5°C different from each other and suboptimal thermocycling conditions. As mentioned earlier, competition for resources in the tube can produce an inefficient reaction.

Whether an efficiency for a target is high or low, matching efficiencies between a target and a normalizer is quite important for maintaining data accuracy. For example, an efficiency of 95% for target A and 96% for normalizer B is more desirable than an efficiency of 99% for target A and 92% for normalizer B.

The problem with skewed efficiency

Efficiencies outside the range of 90–110% may artificially skew results and lead to false conclusions, mainly because targets for comparison will have different efficiencies. In addition, inhibition and poor efficiency can affect assay sensitivity, leading to a smaller dynamic range and decreased versatility (Figure 32).

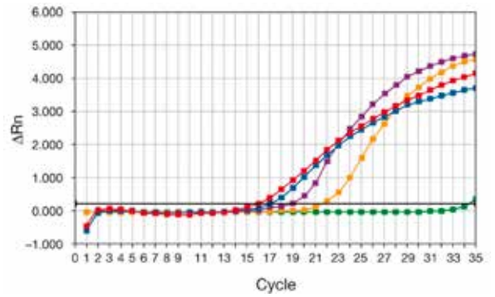


Figure 32. Template dilution series to assess reaction efficiency. Dilutions with earlier C_t values exhibit compressed C_t values and abnormal curve shapes. As the template becomes more dilute, inhibition vanishes and the curves take on the more characteristic exponential phase shape.

Determining if efficiency is skewed

As mentioned earlier, the best method for determining whether a particular assay is inefficient is to generate a standard curve of template diluted over the range of what will be encountered with the unknown samples and look at the efficiency over that range. It should be as close to 100% as possible.

A melting curve or gel showing multiple peaks or products means there is a competition for reaction resources that almost certainly will have an effect on the reaction efficiency.

Resolving poor efficiency or inhibition

Once it has been determined that the reaction is inhibited or is operating with poor efficiency, there are some steps that can be taken to bring the efficiency value back into the desirable range.

1. For inhibition, those wells with the highest concentration of template can be removed and the standard curve reanalyzed. If the efficiency improves back to under 110%, the assay is fine. Just keep in mind that any concentrations removed from the standard curve may not be used during the actual assay.
2. Another solution involves repurifying the template. Remember to allow extra drying time to remove ethanol from ethanol precipitations or to employ additional on-column washes to remove chaotropic salts from silica-based purifications.
3. Poor efficiency is resolved through assay optimization. Sometimes the process can be relatively pain-free, but in other situations, as assay complexity increases, optimization can be laborious.
4. Raising the magnesium concentration as high as 6 mM can improve efficiency in situations where a single product is amplified, but lowering the magnesium may help in cases where competition is occurring.
5. In some circumstances, mainly multiplexing reactions, a primer and probe optimization matrix is necessary. In this application, different ratios or concentrations of forward primer to reverse primer, and sometimes even probe ratios, are tested to find the ideal concentration combination for a given assay. The ideal primer concentration can be anywhere from 100 to 600 nM, while probe concentrations can be between 100 nM and 400 nM.
6. Ensure that the thermal cycling conditions (especially the annealing temperature) are favorable based on the T_m values of the primers, and that the primers are designed to have similar T_m values.

In some cases, issues that appear to be reaction-related may in fact be software-related. Validating and/or optimizing software settings can often bring results back in line with expectations.

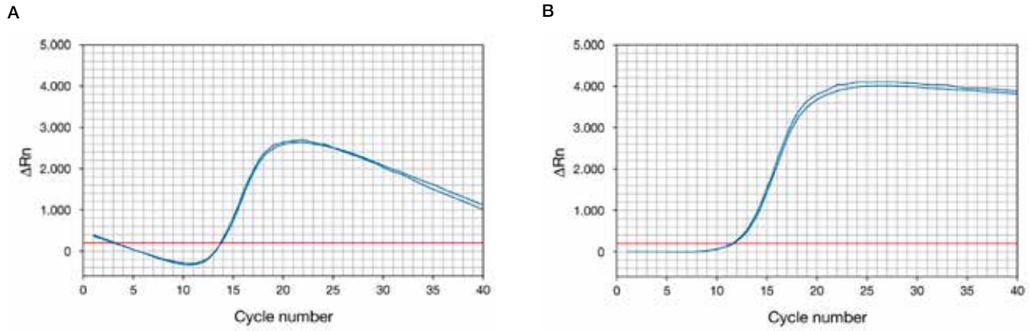


Figure 33. Amplification plots showing (A) an incorrectly set baseline and (B) a correctly set baseline.

Software analysis settings

As mentioned, there are cases in which the instrument analysis may be masking an otherwise successful assay. Analysis settings that play the largest role in data accuracy are:

- Amplification curve baseline linearity
- Baseline range settings
- Threshold
- Reference dyes

The amplification curve baseline linearity is one parameter that can affect results. The instrument software is usually good at automatically setting the baseline within the flat portion of the curve. However, in cases where a very early C_t is observed, such as in cases where 18S rRNA is used as a normalizer, the baseline can mistakenly be placed to include a region that is no longer flat. Figure 33 shows the same plot with different baseline settings. Figure 33A shows a plot with a baseline that spans cycles 1 through 14, which is too wide because fluorescence is detected as early as cycle 10. The result is a curve that dips down and pushes the C_t later. Figure 33B shows the baseline reset to the linear range of cycles 2 through 8, which returns the curve and C_t to their accurate locations.

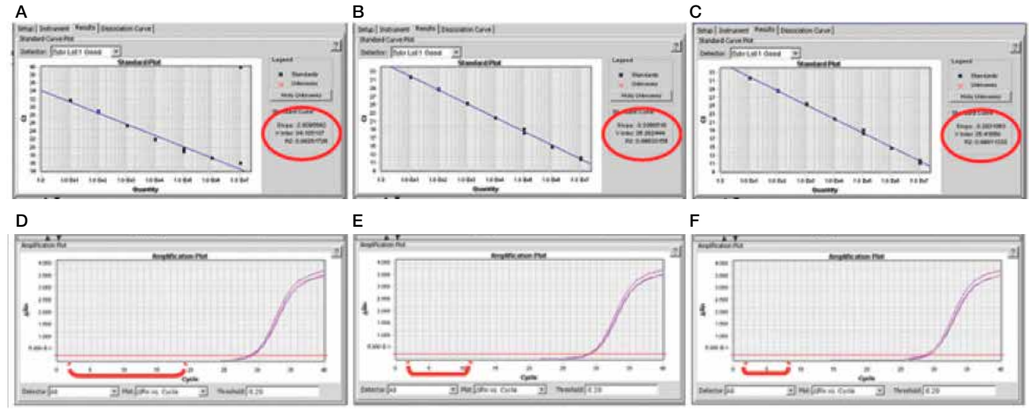


Figure 34. Comparison of baseline setting methods to achieve acceptable reaction efficiencies. (A and **D)** The baseline in these plots is manually set very wide and the slope of the standard curve is poor: only –2.81, which is well outside the preferred range of –3.58 to –3.10 (corresponding to 90–110% efficiency). **(C** and **F)** The same curves as in panels A and D, but with a baseline that the instrument automatically chose. The slope is now inside the ideal window and the assay is now validated across this dynamic range. **(B** and **E)** These plots were manually adjusted to have the baseline incorporate 4 to 5 additional cycles, and the slope improved even further to nearly 100% efficiency.

Baseline range settings are not often considered but can have an effect on the reaction efficiency. As shown in Figure 34, the instrument default settings are more often than not acceptable for a given assay. However, manual adjustment can sometimes improve results even further.

The threshold (the level of fluorescence deemed above background and used to determine C_t) is another parameter set automatically by the software, but one that may also be manually adjusted (Figure 35).

When evaluating more than one kit or chemistry on the same run, the following situation can often occur. The software will automatically select a threshold that is ideal for the curves with the higher plateau (the blue plots in Figure 35). This would bias the C_t values for the red plots in that data set because the ideal threshold is much lower. Therefore, each data set should be studied independently so that the ideal threshold may be selected for each situation.

Again, while the default settings are often very appropriate, the threshold can be manually dragged to the middle of the exponential phase for greater accuracy if needed.

Standard curve dynamic range validation determines what template concentrations are acceptable in a given assay. Concentrations from the high and/or low ends of a standard curve can also be removed to improve the efficiency of a reaction, as long as those concentrations are never employed during the actual assay (Figure 36).

In general, it is OK to push the limits of detection from very high template to very low template, knowing that terminal data points can always be removed if efficiency is compromised.

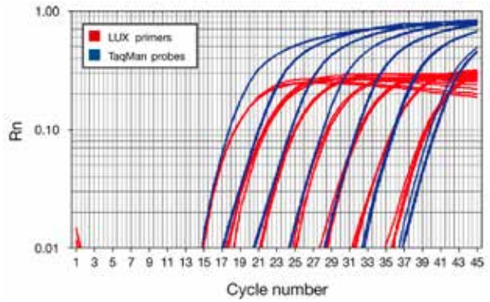


Figure 35. A log plot screen shot showing an example of two sets of curves with differing plateau heights and therefore different exponential phases. The most accurate portion of an amplification curve for C_t determination is directly in the middle of the exponential phase when viewing the log plot. The ideal threshold setting is sometimes unique for each set of data.

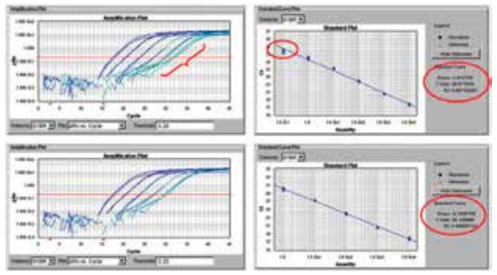
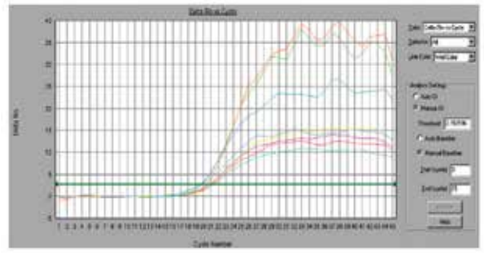


Figure 36. Improvement of standard curve slope achieved by excluding outlier data. Amplification was performed on a dilution series over 5 orders of magnitude. Across this range, the slope is only –2.9, which is outside the desired efficiency window. The curves represented by the lower panels have had the highest dilutions omitted from the standard curve, and the efficiency reanalyzed. The slope has improved to –3.1, which is considered efficient enough to validate the assay.

A



B

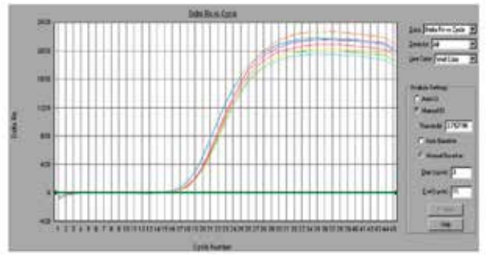


Figure 37. Correction for reference dye. (A) A high level of a passive reference dye such as ROX™ dye can lead to poor target signal returns. **(B)** Once the signal from ROX dye is removed from the analysis, the target signals fall within the expected range.

Employing reference dyes such as ROX dye and fluorescein is a powerful method of insulating against some instrument- and user-related errors. It has become so common that this “behind-the-scenes” factor is often forgotten or assumed not to have a negative impact on a reaction. However, it is important to understand the relationship between the instrument software and the dye itself.

For instruments that employ a reference dye, the software reports the fluorescence signal at Rn (normalized reporter value), which is the reporter dye signal divided by the reference dye signal. Therefore, if the level of ROX dye, for example, is too high, it can result in very poor target signal returns, which manifest as jagged and inconsistent amplification plots (Figure 37).

Visually, it seems as if the reaction failed and much optimization is necessary, but where to start? If the ROX dye channel is switched off as the normalizer and data are reanalyzed, it can be seen that the data are actually just fine (Figure 37B); it is a reference dye normalization issue. Along the same lines, if a mix was used that did not have ROX dye, it is important to make sure that the reference dye option in the software is set to “None” for the analysis.

As mentioned, the default instrument software settings are fine in most situations. However, verification of these settings can increase confidence in data accuracy. Ensure that the baseline chosen by the software is only in the flat range of the amplification curve, and increase or decrease the baseline range as necessary. Look at the amplification curves in the log plot view and verify that the threshold is set near the middle of the exponential phase of the curve. Adjust the y-axis to be appropriately scaled for the fluorescence plateau intensity, and keep in mind that outliers and whole dilution sets may be removed from the standard curve to improve efficiency and R^2 values (as long as those dilutions will not be used when evaluating the “unknown” samples).

Lastly, keep in mind that threshold settings need to be identical for a particular assay when comparing across a data set.

Troubleshooting is an inevitable aspect of real-time PCR assay validation and employment. However, by categorizing and understanding the key issues, this can be a relatively simple process:

- Ensure that primer-dimers are not contributing to signal or poor reaction efficiency
- Take the steps necessary to maintain primer and probe stability
- Make standard curve validation the final step in the reaction assessment process
- Understand that efficiencies below 90% will be addressed very differently from values above 110%
- Verify and adjust instrument analysis settings as necessary

No amplification

One final problem that may occur is a complete lack of amplification using a given assay. After verifying that all the steps above have been addressed, other causes of no amplification include: low expression, problems with reverse transcription, or problems with assay design.

Problems with low expression

The common cDNA input for gene expression analysis is 1 to 100 ng, but if the gene of interest is of low abundance in the sample, more may be needed. Test a range of input, or ideally, run against a positive control sample to confirm that the assay is functioning properly. If unsure about the expected level of expression, check the literature or the NCBI Unigene database for expression data that can give an estimate of expected levels in different tissues.

Problems with reverse transcription

Related to low expression, if the gene of interest is of low abundance in the sample, it may be necessary to increase the sensitivity of the assay. Check that the real-time PCR reaction is not overloaded with cDNA (maximum load is 20% v/v), as this can introduce inhibitors into the reaction and thus reduce the efficiency. Also check the type of reverse transcriptase and primers being used. Random primers typically yield more cDNA than oligo(dT)-based methods. Additionally, some enzymes, such as Invitrogen™ SuperScript™ III Reverse Transcriptase, have been engineered to be more thermostable, which will also increase yield. Check the reaction components to see if any of these elements can be optimized to improve amplification.

Problems with assay design

If no amplification is observed in the assay, it is possible that a primer or probe is not designed to the right target. Check sequence databases such as NCBI for variants of the gene of interest. It is possible that the assay is designed to only one variant, which may not be expressed in the samples being studied. Also check where the primers and probe fall along the target sequence. Is it in a coding region or intron? Assays targeting the 5′ UTR of a gene, for example, will not detect an exogenous gene target from a transfected cell (since the UTR region would not have been included in the plasmid for transfection). Likewise, an assay sitting within an intron sequence will not amplify a cDNA sample.

5.2 Frequently asked questions

Q: How many copies are in a given amount of human gDNA?

A: 1 genome copy = 3 x 10^9 bp
1 bp = 618 g/mol
1 genome copy = (3 x 10^9 bp) x (618 g/mol/bp)
= 1.85 x 10^{12} g/mol = (1.85 x 10^{12} g/mol) x (1 mole/6.02 x 10^{23} [Avogadro's number])
= 3.08 x 10^{-12} g
Each somatic cell has 6.16 pg of DNA (sperm and egg cells have 3.08 pg). There is one copy of every non-repeated sequence per 3.08 pg of human DNA. Therefore, 100 ng of gDNA would have: (100,000 pg of DNA)/3.08 pg = ~33,000 copies; 1 ng of DNA has 330 copies.

Q: Why do I have to be concerned about the efficiency of my real-time PCR assay?

A: If you want to compare the expression levels of two genes (for example, in cases where a normalizer gene is employed), you need to know something about the efficiencies of the PCR to confirm that the C_t values you are observing are not being influenced by contaminants in the PCR reagents or are not arising from a poorly optimized assay.

Q: I have found that my more concentrated template samples give me less efficient amplification curves: Dilute sample gives a slope of –3.4 and an R^2 value of 0.99; concentrated sample gives a slope of –2.5 and an R^2 value 0.84. Why?

A: Something in your sample is inhibiting the PCR. The reason you get better efficiency with the more diluted samples is because the inhibitor (salt or some other component) has been diluted below its inhibitory effect. Here are some references that explain this:

- Ramakers C, Ruijter JM, Deprez RH et al. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* 339:62–66.
- Liu W and Saint DA (2002) A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Anal Biochem* 302:52–59.
- Bar T, Ståhlberg A, Muszta A et al. (2003) Kinetic outlier detection (KOD) in real-time PCR. *Nucleic Acids Res* 31:e105.

Q: Can I compare C_t values of PCR reactions with different efficiencies?

A: You should not compare C_t values of PCR reactions with different efficiencies, because the $\Delta\Delta C_t$ calculation method works on the assumption that PCR efficiencies are comparable. This is why you should optimize your system before trying to quantify unknown samples. The standard curve method of comparative quantification with efficiency correction can be employed.

Q: What are quenchers, and why are they used in real-time PCR?

A: Quenchers are moieties attached to primers or probes so that they can quench the emission from a fluorophore that is also attached to that primer or probe. Quenchers are generally used in probe-based assays to extinguish or change the wavelength of the fluorescence emitted by the fluorophore when both are attached to the same oligonucleotide. They usually do this by fluorescence resonance energy transfer (FRET). When the fluorophore gets excited, it passes on the energy to the quencher, which emits the light at a different (higher) wavelength. Common quenchers are TAMRA dye, or nonfluorescent quenchers such as MGB-NFQ, QSY, or Biosearch Technologies™ Black Hole Quencher™ dyes.

Q: When would I use one-step as opposed to two-step qRT-PCR?

A: Two-step qRT-PCR is popular and useful for detecting multiple messages from a single RNA sample. It also allows the archiving of cDNA for further analysis. However, one-step qRT-PCR is easier to use when processing large numbers of samples and helps minimize carryover contamination, since tubes do not need to be opened between cDNA synthesis and amplification. Since the entire cDNA sample is amplified, one-step qRT-PCR can provide greater sensitivity, down to 0.1 pg total RNA.

Q: What is MGB-NFQ? What is the benefit of it as a quencher?

A: MGB-NFQ stands for minor groove binder–nonfluorescent quencher. The MGB moiety increases the T_m of the probe, thus allowing for the design of shorter, more specific probes. In general, the TaqMan MGB probes exhibit great differences in T_m values between matched and mismatched probes, which enables more accurate allelic discrimination and makes for a more sensitive real-time PCR assay.

Advanced topics

6

Contents

6.1 Digital PCR

Digital PCR compared to traditional real-time PCR

Digital PCR employs the same primer sets, fluorescent labels, and enzymatic reagents as traditional real-time PCR. TaqMan Assays are ideally suited to perform digital PCR; however, SYBR Green dye has demonstrated compatibility. The primary difference between real-time PCR and digital PCR is that in digital PCR, a sample is partitioned into thousands of individual PCR reactions—in essence generating a limiting dilution. Other key differences are detailed in Table 1. In contrast to real-time PCR, digital PCR offers a highly precise and sensitive approach without the need for a reference or standard curve. These key attributes are driven by the number of partitions and volume sampled in digital PCR. The Applied Biosystems™ QuantStudio™ 3D Digital PCR System leverages a chip-based technology, optimally partitioning a standard PCR reaction mix into 20,000 individual PCR reactions (Figure 38). Upfront sample dilution ensures that a portion of these partitions contain the target molecule, while other partitions do not, leading to positive and negative reactions, respectively. Following amplification on a dual flat-block thermal cycler, the fraction of negative reactions is used to generate an absolute count of the number of target molecules in the sample, all without reference to standards or controls. Figure 39 shows the basic procedure used in digital PCR.

Target quantification in digital PCR

Quantification by digital PCR is achieved using fairly simple statistical analysis. Since each reaction is expected to contain zero, one, or a few molecules, the ratio of positive and negative signals will follow a classical Poisson distribution. For example, for a viral DNA sample in a digital PCR reagent mixture that contains 20,000 copies of the viral target, when the mixture is split into 20,000 partitions, mathematically the expectation is to have approximately 1 copy in each reaction. Of course, by chance, there would be a significant number of reactions that contain zero, two, or more than two copies—the probability of these outcomes is described by the Poisson model.

Figure 40 shows the Poisson distribution model. Continuing with this example, if 20% of the 20,000 digital PCR reactions gave a negative signal, the number of target copies in each can be identified by finding 20% on the x-axis and identifying the corresponding average copies per reaction based on the dashed line on the graph. In this example, the result is 1.59 copies/reaction. Since the calculation is based on a percentage, the answer will be the same regardless of the number of reactions. The difference, however, is that with more reactions, the confidence interval is narrower so the statistical reliability of the data is improved.

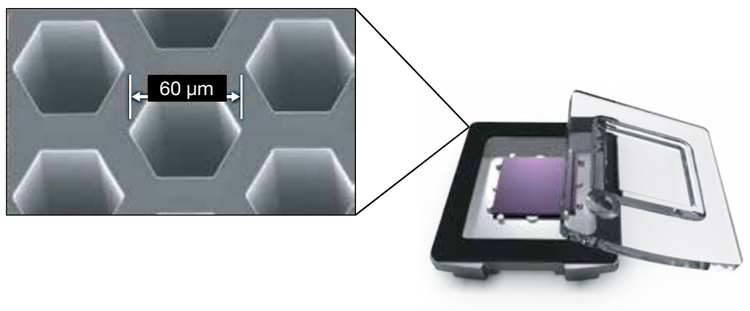


Figure 38. QuantStudio 3D Digital PCR 20K Chip. Each chip is designed with 48 subarrays x 64 through-holes/subarray. Hydrophilic and hydrophobic coatings on plates enable reagents to stay in the bottomless through-holes via capillary action.

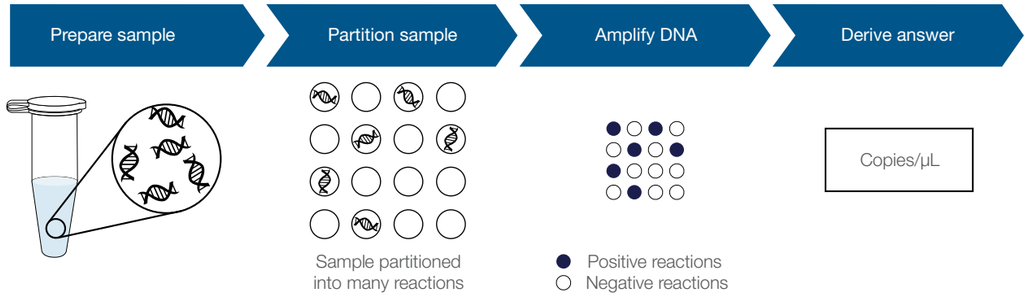


Figure 39. Digital PCR employs a simple workflow and uses familiar techniques.

Table 1. Comparison of conventional real-time PCR to digital PCR.

	Conventional real-time PCR	Digital PCR
Output from experiment	C_t, ΔC_t, or $\Delta\Delta C_t$	Copies per µL
Quantification	Relative quantification	Absolute quantification
Results can be affected by the:	Detection chemistry (e.g., TaqMan Assays or SYBR Green dyes) Real-time PCR instrument used Amplification efficiency of PCR primers/probe	None of these factors

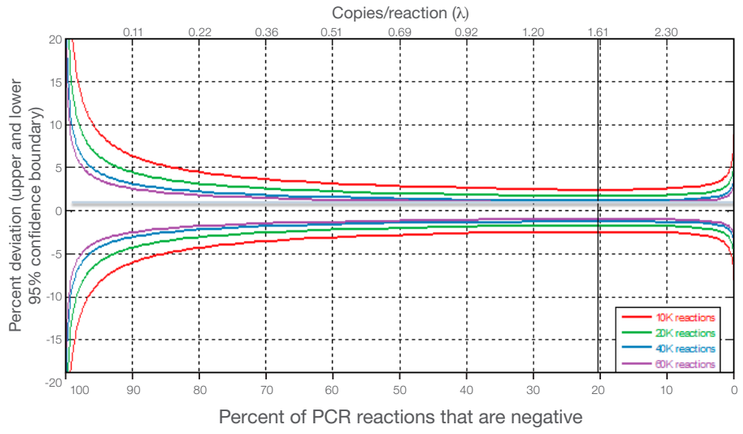


Figure 40. Poisson equation used to calculate target quantity from digital PCR data.

6.2 Digital PCR attributes

Detection of low levels of pathogen

Digital PCR extends the performance of TaqMan Assays by enabling additional attributes that go beyond the limits of real-time PCR. These attributes represent three main categories—increased precision, increased sensitivity, and increased specificity, with the ability to perform absolute quantification without a standard curve.

How does digital PCR manage this extra sensitivity, specificity, and precision?

Sensitivity is driven by total volume interrogated. It's actually not that different from real-time PCR; however, digital PCR yields a statistical perspective via its large number of replicates for the target, giving greater visibility to whether or not the target of interest is detectable. Imagine a container full of balls (Figure 41). The chances of capturing a particular ball of interest from the container are increased with a greater number of balls taken out of the container.

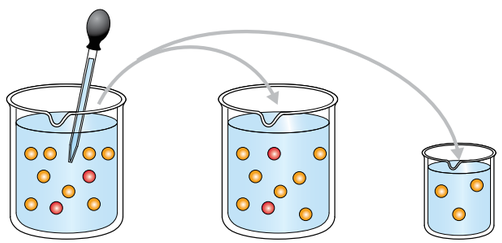


Figure 41. The Poisson principle assumes an appropriate volume of the total pool is sampled. Increasing the amount sampled from the total pool increases the ability to accurately determine the number of targets.

Specificity is driven by the assay and number of replicates run. By individualizing the reaction, digital PCR enables extending the performance of current TaqMan real-time PCR assays in order to drive additional specificity. For example, a sample containing 99 wild type molecules and 1 mutated molecule equates to the mutation being present at 1 in 100 or 1%. Using Applied Biosystems™ TaqMan™ SNP Genotyping Assays in standard real-time PCR mode, the single mutant is lost in a sea of wild type copies (Figure 42A). By first partitioning the sample, competing wild type sequences in any reaction containing a mutant are reduced, effectively decreasing background noise. If sufficient partitions are used, the reaction wells reach a point where the wild type signal no longer overwhelms the mutant signal. In the example in Figure 42B, dividing the sample into twenty digital partitions reduces the sample complexity within each partition to 1 in 5 or 20%—theoretically a 20-fold improvement compared to the starting sample.

Precision is driven by the number of replicates that are run. Increasing replicates increases the statistical significance of the answer, thereby giving more confidence that the value determined represents the actual target quantity in the sample. For maximum precision, the percentage of negative reactions should be targeted between 5% and 80%.

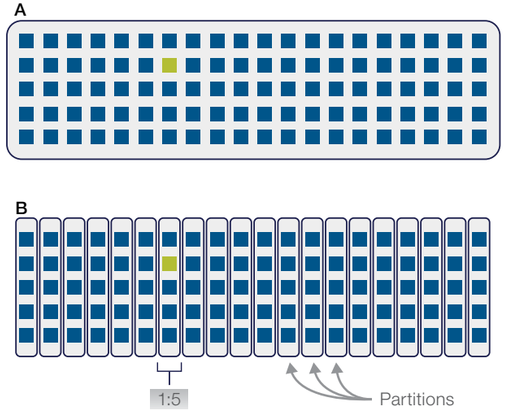


Figure 42. By individualizing the reaction, digital PCR extends the performance of current TaqMan real-time PCR assays in order to drive additional specificity.

6.3 Digital PCR applications

Precise copy number variation (CNV)

CNV is defined as a modification in the genome where the number of copies of a gDNA sequence differs from a reference or standard. Genomic alterations such as insertions, deletions, inversions, or translocations can lead to biallelic or multiallelic CNVs. CNVs are linked to susceptibility or resistance to disease, and thus are an important area for detailed study. Many methods of CNV detection exist today, including fluorescent *in situ* hybridization, comparative genomic hybridization, array comparative genomic hybridization, real-time PCR, and next-generation sequencing.

Despite advances in some of these technologies, in many cases, measurements are not sufficiently precise for determining copy number differences where the ratios between the target and reference are very small.

Digital PCR, a technology capable of highly precise measurements, enables low-percent copy number differences to be detected and accurately quantified.

A representative panel of nine gDNA samples, procured from the Coriell repository, was analyzed using the QuantStudio 3D system and a standard Applied Biosystems™ TaqMan™ Copy Number Assay specific to the *CCL3L1* genetic locus found on the long arm of human chromosome 17, 6, or 8. Replicate measurements indicated that the samples represent copy number variations from 0 to 8 copies per genome (Figure 43A). A statistically significant difference between samples containing 7 and 8 copies was clearly discernable as a result of the high degree of precision achieved, confirming that digital PCR can differentiate less than a 1.2-fold difference (Figure 43B).

A

Sample	Number of replicates	Expected copy number	Detected copy number (mean)	Standard deviation	CV (%)
NA17245	6	0	0.08	0.06	N/A
NA17251	6	1	0.98	0.02	2.21
NA17258	6	2	1.96	0.05	2.47
NA17132	6	3	2.98	0.06	1.85
NA19194	8	4	4.00	0.05	1.22
NA18507	8	5	5.11	0.13	2.50
NA17110	8	6	5.91	0.12	2.07
NA17202	8	7	7.02	0.07	1.02
NA18854	8	8	7.95	0.20	2.55

B

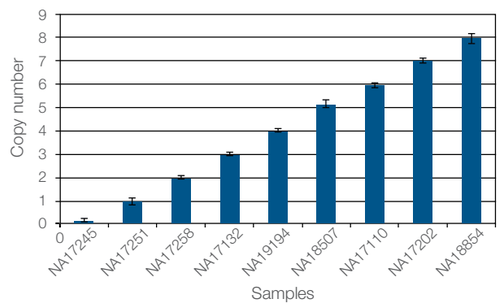


Figure 43. Precision demonstrated for copy number analysis of the *CCL3L1* genetic locus on chromosome 17. (A) Copy number was measured across nine DNA samples. The CV (column 6) was below 2.6% for each set of replicates, demonstrating a high degree of measurement reproducibility within each replicate group. **(B)** As demonstrated by nonoverlapping error bars, the achieved measurement precision enables statistical discernment of the *CCL3L1* copy number in samples containing 7 and 8 copies.

Rare-allele detection

Rare-mutation detection has great implications in areas such as cancer research because the accumulation of mutations in crucial regulatory genes, such as oncogenes or tumor suppressor genes, is an important aspect of tumorigenesis. Acquisition of these mutations in a tiny subset of somatic cells can be sufficient for cancer initiation or progression.

Since these mutations are so rare, they require an assay that delivers high signal-to-noise ratios and low rates of false positives to false negatives.

Common SNP genotyping technologies, such as capillary electrophoresis sequencing and real-time PCR, are most effective at detecting mutant cells with a prevalence no lower than about 20% (or approximately 1 in 5 cells). By combining real-time PCR chemistries, such as TaqMan Assays, with digital PCR methodology, researchers are now able to detect mutant cell prevalence down to 1%—and below (Figure 44).

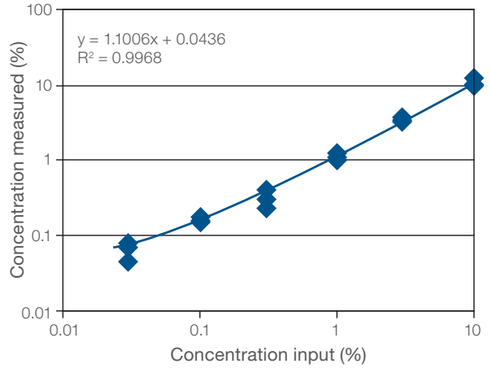


Figure 44. Rare-allele measurement using spike-recovery method. Differing amounts of DNA from three different oncogenic *KRAS* alleles were spiked into a constant amount of normal DNA. Note the excellent correlation between input concentration and measured concentration; the linear slope indicates that the amounts of mutant allele were accurately measured.

Digital PCR works by partitioning a sample into many individual reactions prior to amplification, reducing competing wild type sequences in any reaction containing a mutation and effectively decreasing background noise. If sufficient partitions are used, the reaction wells reach a point where the wild type signal no longer overwhelms the mutant signal. Because each data point is generated digitally, the total count of each allele, mutant and wild type, can be calculated and a ratio determined (Figure 45).

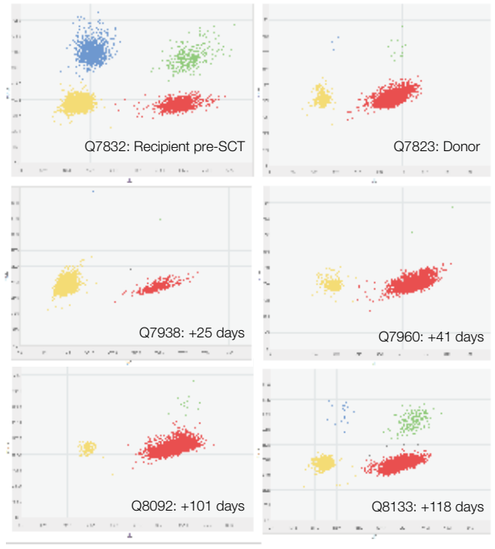


Figure 45. Allelic chimerism in bone marrow transplant samples. Two alternate alleles that differentiated a bone marrow donor from a recipient were chosen. Samples were collected pre–stem cell transplant (pre-SCT) and at the indicated times after transplant. Note the recipient's allele starts to reappear after 101 days, and is obvious by 118 days, indicating a relapse.

Absolute quantification of next-generation sequencing libraries

Next-generation sequencing libraries can be quantified with minimal sample handling and without the need to generate a standard curve using digital PCR. This method enables accurate and precise library quantification, a critical step in both the Ion Torrent™ and other next-generation sequencing workflows, allowing for maximizing sequencing yields downstream. To achieve this high degree of precision, a TaqMan Assay, designed to span both the forward and reverse adapters specific to each library, is available.

This approach limits quantification to library constructs that contain both adapter sequences. Ultimately, using digital PCR to quantify next-generation sequencing libraries decreases overall sequencing costs by ensuring an accurate quantification upfront, minimizing the need to rerun or repeat sequencing of samples. For more information about this application, go to **thermofisher.com/dpcrngs**

Absolute quantification of nucleic acid standards

Accurate genetic measurements often require comparison to reference samples and assay standards. Standardization of references is especially important in the field of metrology. For many organisms or applications, there is often no suitable reference sample available. Generation of reference standards using conventional real-time PCR requires consideration of how the reference sample will initially be calibrated, its long-term stability, and whether there is sufficient reference material for completion of all future studies. In addition, the lack of broadly adopted standards impacts comparison between laboratories.

Digital PCR does not rely on a reference sample or assay standard; it can be used for absolute quantification, measuring the exact copy number of a nucleic acid target of interest (Figure 46). This capability is especially useful for calibrating reference samples and assay standards when none exist. Through direct copy number determination, digitally measured assay standards can enable laboratories to compare results, with the assurance that measurements are based on the same absolute baseline.

Low-level fold change of gene expression

Real-time PCR is commonly used to detect differential gene expression; however, this approach is generally limited to detecting changes that vary by 2-fold or more. For some studies, detection of expression changes less than 2-fold may be required. Furthermore, it is often necessary to express differential gene expression with respect to a reference gene, such as a housekeeping gene like actin.

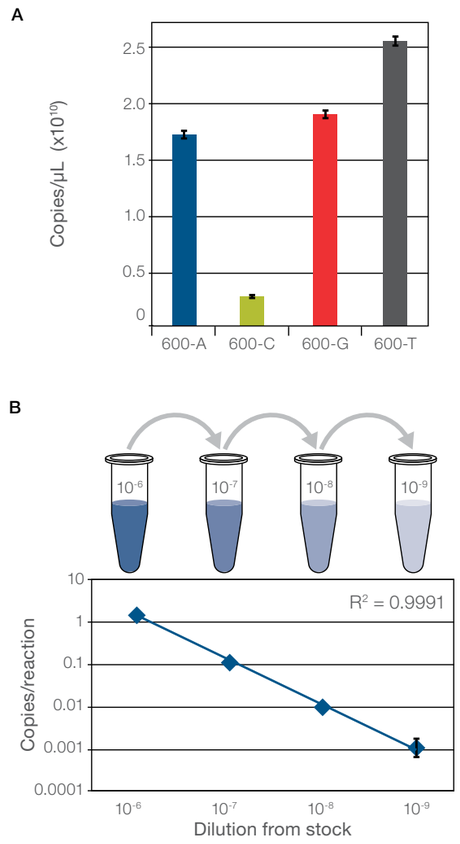


Figure 46. Digital PCR precisely and accurately quantifies standards without the use of a standard curve. (A) Four standards were measured in duplicate, and results determined in absolute copies per microliter by digital PCR. The tight error bars demonstrate very high precision in the measurement of each sample. **(B)** For sample 600-T, an additional 10-fold dilution series over 4 orders of magnitude was constructed. Copies per reaction for each dilution were calculated and demonstrate excellent correlation (0.9991), with extremely tight precision for each dilution.

With the ability to achieve highly precise measurements of ±10% or better, digital PCR is capable of resolving changes of 2-fold or less (Figure 47).

In addition, the ability of digital PCR to determine absolute quantification of a transcript obviates the need for a reference gene. Like real-time PCR, digital PCR requires the conversion of RNA to cDNA. Offering the efficiency of conversion that is so important to experimental sensitivity, the Applied Biosystems™ High-Capacity cDNA Reverse Transcription Kit seamlessly integrates into a digital PCR gene expression workflow.

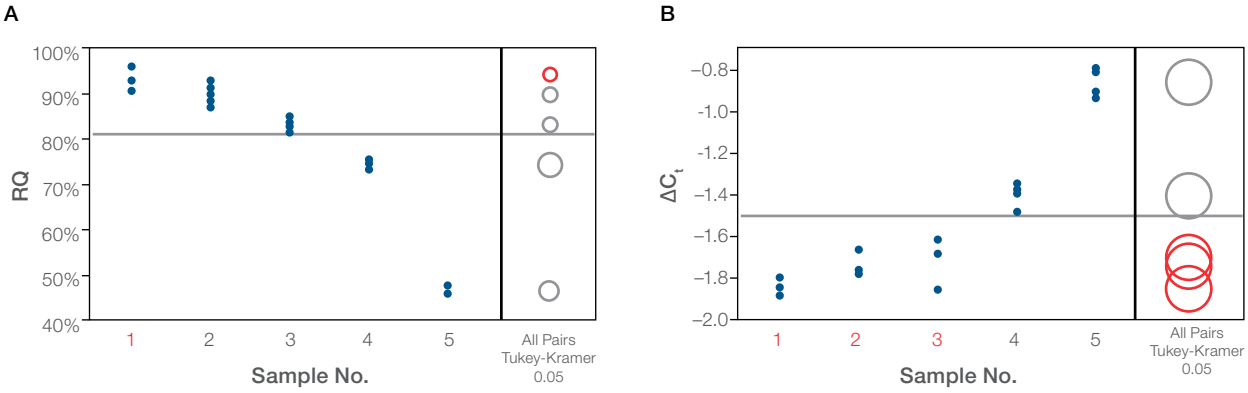


Figure 47. Quantitation precision comparison between (A) digital PCR and (B) real-time PCR. Samples 1 through 5 are mixtures of synthetic miRNAs hsa-miR-19b and hsa-miR-92 at different ratios: sample 1, 100%; sample 2, 95%; sample 3, 90%; sample 4, 75%; sample 5, 50%. After reverse transcription, cDNA was measured by real-time PCR and digital PCR on the QuantStudio 3D system. ΔC_t values of real-time PCR between hsa-miR-19b and hsa-miR-92 were reported for each sample. The relative quantitation from digital PCR results is reported in percentiles for each sample. Digital PCR with the QuantStudio 3D system is able to discriminate a 5% difference between samples 1 and 2 (indicated by nonoverlapping circles by Tukey-Kramer HSD test), while real-time real-time PCR was not able to discriminate even a 10% difference between samples 1 and 3. Tukey-Kramer HSD test was done within JMP software with experimental replicates.

6.4 Beginning a digital PCR experiment

To perform a digital PCR experiment, the sample must be diluted such that each reaction contains one or zero molecules. First establish the starting concentration using a spectrophotometer and convert the concentration in ng/µL into copies/µL. Next, use this value to calculate the volume of sample needed to target 20,000 copies/chip for each sample. If the target copy number per genome of the samples is known, dilute the samples so that, when loaded on a QuantStudio 3D Digital PCR 20K Chip, each through-hole reaction will contain approximately 0.6 to 1.6 copies of the target sequence. For example, assuming 3.3 pg/copy of a given gene is present per genome and a 865 pL reaction well volume, the stock gDNA in a given sample would be diluted down to 600 copies/µL or 1.98 ng/µL in the final reaction to give 0.6 copies per reaction well. Refer to the QuantStudio 3D Digital PCR System product manual to learn how to determine target copy number per genome.

Each application has its own set of factors to consider when setting up a digital PCR experiment. Please refer to the QuantStudio 3D Digital PCR System Experimental Design Guide for more detailed application-specific instructions. This can be found at **thermofisher.com/quantstudio3d** or in the community on the digital PCR forum at **thermofisher.com/dpcrcommunity**

applied
biosystems

Find out more at **thermofisher.com/qpcr**

ThermoFisher
SCIENTIFIC